

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGAGAATTAGATCGATGGGAAAAAAT
TCGGTTAAGGCCAGGGGGAAAGAAGAAGTACAAGCTAAAGCACATCGTATGGGCAA
GCAGGGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGC
TGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAGGAGCT
TCGATCACTATACAACACAGTAGCAACCCTCTATTGTGTGCACCAGCGGATCGAGA
TCAAGGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAGTCCAAG
AAGAAGGCCCAGCAGGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAA
TTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTA
GAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTG
ATACCCATGTTTTTCAGCATTATCAGAAGGAGCCACCCACAGGACCTGAACACGAT
GTTGAACACCGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCA
ATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCA
CCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCT
TCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAGATCT
ACAAGAGGTGGATAATCCTGGGATTGAACAAGATCGTGAGGATGTATAGCCCTACC
AGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCG
GTTCTATAAACTCTAAGAGCTGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGA
CAGAAACCTTGTTGGTCCAAAATGCGAACCAGATTGTAAGACCATCCTGAAGGCT
CTCGGCCCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGG
ACCCGGCCATAAGGCAAGAGTTTTGGCCGAGGCGATGAGCCAGGTGACGAACTCGG
CGACCATAATGATGCAGAGAGGCAACTTCCGGAACCAGCGGAAGATCGTCAAGTGC
TTCAATTGTGGCAAAGAAGGGCACACCGCCAGGAACTGCCGGGCCCCCGGAAGAA
GGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACTGAGAGAC

FIG. 1

AGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTT
CTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGT
AGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACCTGTATCCTT
TAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAGTAAGGATCGGGG
GGCAACTCAAGGAAGCGCTGCTCGATACAGGAGCAGATGATACAGTATTAGAAGAA
ATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGGATCGGGGGCTTCAT
CAAGGTGAGGCAGTACGACCAGATACTCATAGAAATCTGTGGACATAAAGCTATAG
GTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACC
CAGATCGGCTGCACCTTGAACCTCCCCATCAGCCCTATTGAGACGGTGCCCGTGAA
GTTGAAGCCGGGGATGGACGGCCCCAAGGTCAAGCAATGGCCATTGACGAAAGAGA
AGATCAAGGCCTTAGTCGAAATCTGTACAGAGATGGAGAAGGAAGGGAAGATCAGC
AAGATCGGGCCTGAGAACCCCTACAACACTCCAGTCTTCGCAATCAAGAAGAAGGA
CAGTACCAAGTGGAGAAAGCTGGTGGACTTCAGAGAGCTGAACAAGAGAACTCAGG
ACTTCTGGGAAGTTCAGCTGGGCATCCACATCCCGCTGGGTGAAGAAGAAGAAG
TCAGTGACAGTGCTGGATGTGGGTGATGCCTACTTCTCCGTTCCCTTGGACGAGGA
CTTCAGGAAGTACACTGCCTTCACGATACCTAGCATCAACAACGAGACACCAGGCA
TCCGCTACCAGTACAACGTGCTGCCACAGGGATGGAAGGGATCACCAGCCATCTTT
CAAAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAAACCCAGACATCGT
GATCTATCAGTACATGGACGACCTCTACGTAGGAAGTGACCTGGAGATCGGGCAGC
ACAGGACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGGTGGGGACTGACCACA
CCAGACAAGAAGCACCAGAAGGAACCTCCCTTCCTGTGGATGGGCTACGAACTGCA
TCCTGACAAGTGGACAGTGCAGCCCATCGTGCTGCCTGAGAAGGACAGCTGGACTG
TGAACGACATACAGAAGCTCGTGGGCAAGTTGAACTGGGCAAGCCAGATCTACCCA
GGCATCAAAGTTAGGCAGCTGTGCAAGCTGCTTCGAGGAACCAAGGCACTGACAGA

AGTGATCCCACTGACAGAGGAAGCAGAGCTAGAACTGGCAGAGAACCGAGAGATCC
TGAAGGAGCCAGTACATGGAGTGTACTACGACCCAAGCAAGGACCTGATCGCAGAG
ATCCAGAAGCAGGGGCAAGGCCAATGGACCTACCAAATCTACCAGGAGCCCTTCAA
GAACCTGAAGACAGGCAAGTACGCAAGGATGAGGGGTGCCACACCAACGATGTGA
AGCAGCTGACAGAGGCAGTGCAGAAGATCACCACAGAGAGCATCGTGATCTGGGGC
AAGACTCCCAAGTTCAAGCTGCCCATACAGAAGGAGACATGGGAGACATGGTGGAC
CGAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGAGTTCGTGAACACCCCTCCCT
TGGTGAAACTGTGGTATCAGCTGGAGAAGGAACCCATCGTGGGAGCAGAGACCTTC
TACGTGGATGGGGCAGCCAACAGGGAGACCAAGCTGGGCAAGGCAGGCTACGTGAC
CAACCGAGGACGACAGAAAGTGGTGACCCTGACTGACACCACCAACCAGAAGACTG
AGCTGCAAGCCATCTACCTAGCTCTGCAAGACAGCGGACTGGAAGTGAACATCGTG
ACAGACTCACAGTACGCACTGGGCATCATCCAAGCACCAACCAGACCAATCCGAGTC
AGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTAACCTGG
CATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTC
AGTGCTGGGATCCGGAAGGTGCTGTTCCCTGGACGGGATCGATAAGGCCCAAGATGA
ACATGAGAAGTACCACTCCAACCTGGCGCGCTATGGCCAGCGACTTCAACCTGCCAC
CTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAA
GCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAGCTGGACTGCACGCA
CCTGGAGGGGAAGGTGATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAG
CAGAAGTTATCCCTGCTGAAACTGGGCAGGAAACAGCATATTTTCTTTTAAATTA
GCAGGAAGATGGCCAGTAAAAACAATACACACGGACAACGGAAGCAACTTCACTGG
TGCTACGGTTAAGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTC
CCTACAATCCCCAATCGCAAGGAGTCGTGGAGAGCATGAACAAGGAGCTGAAGAAG
ATCATCGGACAAGTGAGGGATCAGGCTGAGCACCTGAAGACAGCAGTGCAGATGGC

AGTGTTTCATCCACAACCTTCAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGG
AAAGGATCGTGGACATCATCGCCACCGACATCCAAACCAAGGAGCTGCAGAAGCAG
ATCACCAAGATCCAGAACTTCCGGGTGTACTACCGCGACAGCCGCAACCCACTGTG
GAAGGGACCAGCAAAGCTCCTCTGGAAGGGAGAGGGGGCAGTGGTGATCCAGGACA
ACAGTGACATCAAAGTGGTGCCAAGGCGCAAGGCCAAGATCATCCGCGACTATGGA
AAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACCT
GGAAGAGCCTGGTGAAGCACCATATG (SEQUENCE ID NO:1)

>wildtype	TGTACAGAGA	TGGAAAAGGA	AGGGAATTTG	TCAAAAAATTG
>mutated #1	TGTACAGAGA	TGGAGAAGGA	AGGGAAGATC	AGCAAGATCG
	+	+ + *** + +	
>wildtype	GGCCTGAAAA	TCCATACAAT	ACTCCAGTAT	TTGCCATAAA
>mutated #41	GGCCTGAGAA	CCCCTACAAC	ACTCCAGTCT	TCGCAATCAA
	+ + + + + + + +		
>wildtype	GAAAAAAGAC	AGTACTAAAT	GGAGAAAATT	AGTAGATTTC
>mutated #81	GAAGAAGGAC	AGTACCAAGT	GGAGAAAGCT	GGTGGACTTC
	+ + + + + + + + +		
>wildtype	AGAGAACTTA	ATAAGAGAAC	TCAAGACTTC	TGGGAAGTTC
>mutated #121	AGAGAGCTGA	ACAAGAGAAC	TCAGGACTTC	TGGGAAGTTC
	+ + + + + + + +		
>wildtype	AATTAGGAAT	ACCACATCCC	GCAGGGTTAA	AAAAGAAAAA
>mutated #161	AGCTGGGCAT	CCCACATCCC	GCTGGGTGTA	AGAAGAAGAA
	+ + + + + + + + + +		
>wildtype	ATCAGTAACA	GTA CTGGATG	TGGGTGATGC	ATATTTTTCA
>mutated #201	GTCAGTGACA	GTGCTGGATG	TGGGTGATGC	CTACTTCTCC
	+ + + + + + + + +		
>wildtype	GTTCCCTTAG	ATGAAGACTT	CAGGAAATAT	ACTGCATTTA
>mutated #241	GTTCCCTTGG	ACGAGGACTT	CAGGAAGTAC	ACTGCCTTCA
	+ + + + + + + +		
>wildtype	CCATACCTAG	TATAAACAAAT	GAGACACCAG	GGATTAGATA
>mutated #281	CGATACCTAG	CATCAACAAC	GAGACACCAG	GCATCCGCTA
	+ + + + + + + + +		
>wildtype	TCAGTACAAT	GTGCTTCCAC	AGGGATGGAA	AGGATCACCA
>mutated #321	CCAGTACAAC	GTGCTGCCAC	AGGGATGGAA	GGGATCACCA
	+ + + + + + + + +		
>wildtype	GCAATATTCC	AAAGTAGCAT	GACAAAAATC	TTAGAGCCTT
>mutated #361	GCCATCTTTC	AAAGCAGCAT	GACCAAGATC	CTGGAGCCCT
	+ + + + + + + + +		
>wildtype	TTAGAAAACA	AAATCCAGAC	ATAGTTATCT	ATCAATACAT
>mutated #401	TCCGCAAGCA	AAACCCAGAC	ATCGTGATCT	ATCAGTACAT
	+ + + + + + + + + +		

FIG. 2

```

>wildtype      GGATGATTTG TATGTAGGAT CTGACTTAGA AATAGGGCAG
>mutated       GGACGACCTC TACGTAGGAA GTGACCTGGA GATCGGGCAG
#441           .....
               *  *  *  *  *  *  *  *

>wildtype      CATAGAACAA AAATAGAGGA GCTGAGACAA CATCTGTTGA
>mutated       CACAGGACCA AGATCGAGGA GCTGAGACAG CATCTGTTGA
#481           .....
               *  *  *  *  *

>wildtype      GGTGGGGACT TACCACACCA GACAAAAAAC ATCAGAAAGA
>mutated       GGTGGGGACT GACCACACCA GACAAGAAGC ACCAGAAGGA
#521           .....
               *  *  *  *

>wildtype      ACCTCCATTG CTTTGGATGG GTTATGAACT CCATCCTGAT
>mutated       ACCTCCCTTC CTGTGGATGG GCTACGAACT GCATCCTGAC
#561           .....
               *  *  *  *

>wildtype      AAATGGACAG TACAGCCTAT AGTGCTGCCA GAAAAAGACA
>mutated       AAGTGGACAG TGCAGCCCAT CGTGCTGCCT GAGAAGGACA
#601           .....
               *  *  *  *

>wildtype      GCTGGACTGT CAATGACATA CAGAAGTTAG TGGGGAAATT
>mutated       GCTGGACTGT GAACGACATA CAGAAGCTCG TGGGCAAGTT
#641           .....
               *  *  *  *

>wildtype      GAATTGGGCA AGTCAGATT ACCCAGGGAT TAAAGTAAGG
>mutated       GAACTGGGCA AGCCAGATCT ACCCAGGCAT CAAAGTTAGG
#681           .....
               *  *  *  *

>wildtype      CAATTATGTA AACTCCTTAG AGGAACCAAA GCACTAACAG
>mutated       CAGCTGTGCA AGCTGCTTCG AGGAACCAAG GCACTGACAG
#721           .....
               *  *  *  *

>wildtype      AAGTAATACC ACTAACAGAA GAAGCAGAGC TAGAACTGGC
>mutated       AAGTGATCCC ACTGACAGAG GAAGCAGAGC TAGAACTGGC
#761           .....
               *  *  *  *

>wildtype      AGAAAACAGA GAGATTCTAA AAGAACCAGT ACATGGAGTG
>mutated       AGAGAACCGA GAGATCCTGA AGGAGCCAGT ACATGGAGTG
#801           .....
               *  *  *  *

>wildtype      TATTATGACC CATCAAAGA CTTAATAGCA GAAATACAGA
>mutated       TACTACGACC CAAGCAAGGA CCTGATCGCA GAGATCCAGA
#841           .....
               *  *  *  *

```

```

>wildtype      AGCAGGGGCA AGGCCAATGG ACATATCAAA TTTATCAAGA
>mutated       AGCAGGGGCA AGGCCAATGG ACCTACCAAA TCTACCAGGA
#881           .....
               *      *      *      *      *

>wildtype      GCCATTAAAA AATCTGAAAA CAGGAAAATA TGCAAGAATG
>mutated       GCCCTTCAAG AACCTGAAGA CAGGCAAGTA CGCAAGGATG
#921           .....
               *      *      *      *      *

>wildtype      AGGGGTGCCC AACTAATGA TGTAACAA TTAACAGAGG
>mutated       AGGGGTGCCC ACACCAACGA TGTGAAGCAG CTGACAGAGG
#961           .....
               *      *      *      *      *

>wildtype      CAGTGCAAAA AATAACCACA GAAAGCATAG TAATATGGGG
>mutated       CAGTGCAGAA GATCACCACA GAGAGCATCG TGATCTGGGG
#1001          .....
               *      *      *      *      *

>wildtype      AAAGACTCCT AAATTTAAAC TGCCCATACA AAAGGAAACA
>mutated       CAAGACTCCC AAGTTCAAGC TGCCCATACA GAAGGAGACA
#1041          .....
               *      *      *      *      *

>wildtype      TGGGAAACAT GGTGGACAGA GTATTGGCAA GCCACCTGGA
>mutated       TGGGAGACAT GGTGGACCGA GTACTGGCAA GCCACCTGGA
#1081          .....
               *      *      *

>wildtype      TTCCTGAGTG GGAGTTTGTG AATACCCCTC CTTTAGTGAA
>mutated       TCCCTGAGTG GGAGTTCGTG AACACCCCTC CTTTGGTGAA
#1121          .....
               *      *      *      *      *

>wildtype      ATTATGGTAC CAGTTAGAGA AAGAACCCAT AGTAGGAGCA
>mutated       ACTGTGGTAT CAGCTGGAGA AGGAACCCAT CGTGGGAGCA
#1161          .....
               *      *      *      *      *

>wildtype      GAAACCTTCT ATGTAGATGG GGCAGCTAAC AGGGAGACTA
>mutated       GAGACCTTCT ACGTGGATGG GGCAGCCAAC AGGGAGACCA
#1201          .....
               *      *      *      *

>wildtype      AATTAGGAAA AGCAGGATAT GTTACTAATA GAGGAAGACA
>mutated       AGCTGGGCAA GGCAGGCTAC GTGACCAACC GAGGACGACA
#1241          .....
               *      *      *      *      *

>wildtype      AAAAGTTGTC ACCCTAACTG ACACAACAAA TCAGAAGACT
>mutated       GAAAGTGGTG ACCCTGACTG ACACCACCAA CCAGAAGACT
#1281          .....

```

```

>wildtype      GAGTTACAAG CAATTTATCT AGCTTTGCAG GATTCGGGAT
>mutated       GAGCTGCAAG CCATCTACCT AGCTCTGCAA GACAGCGGGAC
#1321          .....
               * * * * *

>wildtype      TAGAAGTAAA CATAGTAACA GACTCACAAT ATGCATTAGG
>mutated       TGGAAGTGAA CATCGTGACA GACTCACAGT ACGCACTGGG
#1361          .....
               * * * * *

>wildtype      AATCATTCAA GCACAACCAG ATCAAAGTGA ATCAGAGTTA
>mutated       CATCATCCAA GCACAACCAG ACCAATCCGA GTCAGAGCTG
#1401          .....
               * * * * *

>wildtype      GTCAATCAAA TAATAGAGCA GTTAATAAAA AAGGAAAAGG
>mutated       GTGAACCAGA TCATCGAGCA GCTGATCAAG AAGGAGAAAAG
#1441          .....
               * * * * *

>wildtype      TCTATCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG
>mutated       TGTACCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG
#1481          .....
               * *

>wildtype      AAATGAACAA GTAGATAAAT TAGTCAGTGC TGGGAATCAGG
>mutated       AAATGAACAA GTAGATAAAT TAGTCAGTGC TGGGATCCGG
#1521          .....
               * *

>wildtype      AAAGTACTAT TTTTAGATGG AATAGATAAG GCCCAAGATG
>mutated       AAGGTGCTGT TCCTGGACGG GATCGATAAG GCCCAAGATG
#1561          .....
               * * * * *

>wildtype      AACATGAGAA ATATCACAGT AATTGGAGAG CAATGGCTAG
>mutated       AACATGAGAA GTACCACTCC AACTGGCGCG CTATGGCCAG
#1601          .....
               * * * * *

>wildtype      TGATTTTAAC CTGCCACCTG TAGTAGCAAA AGAAATAGTA
>mutated       CGACTTCAAC CTGCCACCTG TAGTAGCAAA AGAAATAGTA
#1641          .....
               * * *

>wildtype      GCCAGCTGTG ATAAATGTCA GCTAAAAGGA GAAGCCATGC
>mutated       GCCAGCTGTG ATAAATGTCA GCTAAAAGGA GAAGCCATGC
#1681          .....

>wildtype      ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAACTAGA
>mutated       ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAGCTGGA
#1721          .....
               * *

```



```

>wildtype      TTGTACACAT TTAGAAGGAA AAGTTATCCT GGTAGCAGTT
>mutated       CTGCACGCAC CTGGAGGGGA AGGTGATCCT GGTAGCAGTT
#1761          .....
          *  *  *  *  *  *  *  *

>wildtype      CATGTAGCCA GTGGATATAT AGAAGCAGAA GTTATTCCAG
>mutated       CATGTAGCCA GTGGATATAT AGAAGCAGAA GTTATCCCTG
#1801          .....
                      *  *

>wildtype      CAGAAACAGG GCAGGAAACA GCATATTTTC TTTTAAAATT
>mutated       CTGAAACTGG GCAGGAAACA GCATATTTTC TTTTAAAATT
#1841          .....
          *          *

>wildtype      AGCAGGAAGA TGGCCAGTAA AAACAATACA TACAGACAAT
>mutated       AGCAGGAAGA TGGCCAGTAA AAACAATACA CACGGACAAC
#1881          .....
                      *  *  *

>wildtype      GGCAGCAATT TCACCAGTGC TACGGTTAAG GCCGCCTGTT
>mutated       GGAAGCAACT TCACTGGTGC TACGGTTAAG GCCGCCTGTT
#1921          .....
          *          *  *

>wildtype      GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA
>mutated       GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA
#1961          .....

>wildtype      TCCCCAAAGT CAAGGAGTAG TAGAATCTAT GAATAAAGAA
>mutated       TCCCCAATCG CAAGGAGTCG TGGAGAGCAT GAACAAGGAG
#2001          .....
          *  *  *  *  *  *  *

>wildtype      TTAAAGAAAA TTATAGGACA GGTAAGAGAT CAGGCTGAAC
>mutated       CTGAAGAAGA TCATCGGACA AGTGAGGGAT CAGGCTGAGC
#2041          .....
          *  *  *  *  *  *

>wildtype      ATCTTAAGAC AGCAGTACAA ATGGCAGTAT TCATCCACAA
>mutated       ACCTGAAGAC AGCAGTGCAG ATGGCAGTGT TCATCCACAA
#2081          .....
          *  *          *  *

>wildtype      TTTTAAAAGA AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG
>mutated       CTTCAAAAAGA AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG
#2121          .....
          *  *

>wildtype      GAAAGAATAG TAGACATAAT AGCAACAGAC ATACAAACTA
>mutated       GAAAGGATCG TGGACATCAT CGCCACCGAC ATCCAAACCA
#2161          .....
          *  *  *  *  *  *  *

```

```

>wildtype      AAGAATTACA AAAACAAATT ACAAAAATTC AAAATTTTCG
>mutated       AGGAGCTGCA GAAGCAGATC ACCAAGATCC AGAACTTCCG
#2201          .....
               *  *  *  *  *  *  *  *  *  *

>wildtype      GGTTTATTAC AGGGACAGCA GAAATCCACT TTGGAAAGGA
>mutated       GGTGTACTAC CGCGACAGCC GCAACCCACT GTGGAAGGGA
#2241          .....
               *  *  *  *  *  *  *  *

>wildtype      CCAGCAAAGC TCCTCTGGAA AGGTGAAGGG GCAGTAGTAA
>mutated       CCAGCAAAGC TCCTCTGGAA GGGAGAGGGG GCAGTGGTGA
#2281          .....
               *  *  *  *  *  *  *  *

>wildtype      TACAAGATAA TAGTGACATA AAAGTAGTGC CAAGAAGAAA
>mutated       TCCAGGACAA CAGTGACATC AAAGTGGTGC CAAGGCGCAA
#2321          .....
               *  *  *  *  *  *  *  *

>wildtype      AGCAAAGATC ATTAGGGATT ATGGAAAACA GATGGCAGGT
>mutated       GGCCAAGATC ATCCGCGACT ATGGAAAACA GATGGCAGGT
#2361          .....
               *  *  *  *  *  *  *  *

>wildtype      GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTAGAACA
>mutated       GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTAGAACC
#2401          .....
               *  *  *  *  *  *  *  *

>wildtype      TGGAAAAGTT TAGTAAAACA CCATATG
>mutated       TGGAAGAGCC TGGTGAAGCA CCATATG
#2441          .....
               *  *  *  *  *  *  *  *

```

ATGGGCGTGAGAACTCCGTCTTGTGAGGGAAGAAAGCAGATGAATTAG
AAAAAATTAGGCTACGACCCAACGGAAAGAAAAAGTACATGTTGAAGC
ATGTAGTATGGGCAGCAAATGAATTAGATAGATTTGGATTAGCAGAAAG
CCTGTTGGAGAACAAAGAAGGATGTCAAAAAATACTTTCGGTCTTAGCT
CCATTAGTGCCAACAGGCTCAGAAAATTTAAAAAGCCTTTATAATACTG
TCTGCGTCATCTGGTGCATTCACGCAGAAGAGAAAGTGAAACACACTGA
GGAAGCAAAACAGATAGTGCAGAGACACCTAGTGGTGGAAACAGGAAC
CACCGAAACCATGCCGAAGACCTCTCGACCAACAGCACCATCTAGCGGC
AGAGGAGGAAACTACCCAGTACAGCAGATCGGTGGCAACTACGTCCAC
CTGCCACTGTCCCCGAGAACCCTGAACGCTTGGGTCAAGCTGATCGAGG
AGAAGAAGTTCGGAGCAGAAGTAGTGCCAGGATTCCAGGCACTGTCAG
AAGGTTGCACCCCCTACGACATCAACCAGATGCTGAACTGCGTTGGAGA
CCATCAGGCGGCTATGCAGATCATCCGTGACATCATCAACGAGGAGGCT
GCAGATTGGGACTTGCAGCACCCACAACCAGCTCCACAACAAGGACAA
CTTAGGGAGCCGTCAGGATCAGACATCGCAGGAACCACTCCTCAGTTG
ACGAACAGATCCAGTGGATGTACCGTCAGCAGAACCCGATCCCAGTAGG
CAACATCTACCGTCGATGGATCCAGCTGGGTCTGCAGAAATGCGTCCGT
ATGTACAACCCGACCAACATTCTAGATGTAAAACAAGGGCCAAAAGAG
CCATTTTCAGAGCTATGTAGACAGGTTCTACAAAAGTTTAAGAGCAGAAC
AGACAGATGCAGCAGTAAAGAATTGGATGACTCAAACACTGCTGATTCA
AAATGCTAACCCAGATTGCAAGCTAGTGCTGAAGGGGCTGGGTGTGAAT
CCCACCCTAGAAGAAATGCTGACGGCTTGTCAAGGAGTAGGGGGGGCCG
GGACAGAAGGCTAGATTAATGGCAGAAGCCCTGAAAGAGGCCCTCGCA
CCAGTGCCAATCCCTTTTGCAGCAGCCCAACAGAGGGGACCAAGAAAGC
CAATTAAGTGTTGGAATTGTGGGAAAGAGGGGACACTCTGCAAGGCAATG
CAGAGCCCCAAGAAGACAGGGATGCTGGAAATGTGGAAAAATGGACCA
TGTTATGGCCAAATGCCCAGACAGACAGGCGGGTTTTTTAGGCCTTGGT
CCATGGGGAAAGAAGCCCCGCAATTTCCCATGGCTCAAGTGCATCAGG
GGCTGATGCCAACTGCTCCCCAGAGGACCCAGCTGTGGATCTGCTAAA
GAACTACATGCAGTTGGGCAAGCAGCAGAGAGAAAGCAGAGAGAAAG
CAGAGAGAAGCCTTACAAGGAGGTGACAGAGGATTTGCTGCACCTCAAT
TCTCTCTTTGGAGGAGACCAAGTAG

FIG. 3

```

SIV gag -----
#1      .....
      ATGGGCGTGAGAAACTCCGTCTTGTCAGGGAAGAAAGCAG

SIV gag -----
#41     .....
      ATGAATTAGAAAAAATTAGGCTACGACCCAACGGAAAGAA

SIV gag -----
#81     .....
      AAAGTACATGTTGAAGCATGTAGTATGGGCAGCAAATGAA

SIV gag -----
#121    .....
      TTAGATAGATTTGGATTAGCAGAAAGCCTGTTGGAGAACA

SIV gag -----
#161    .....
      AAGAAGGATGTCAAAAAATACTTTCGGTCTTAGCTCCATT

SIV gag -----
#201    .....
      AGTGCCAACAGGCTCAGAAAATTTAAAAAGCCTTTATAAT

SIV gag -----
#241    .....
      ACTGTCTGCGTCATCTGGTGCATTACGCAGAAGAGAAAG

SIV gag -----
SIVgagDX.. -----
#281    .....
      TGAAACACACTGAGGAAGCAAAACAGATAGTGCAGAGACA

SIV gag -----A--A----T----A--A
SIVgagDX..-----C--C----C----G--G
#321    .....
      CCTAGTGGTGGAAACAGGAACMACMGAAACYATGCCRAAR

SIV gag  --AAG-A-----
SIVgagDX..--CTC-C-----
#361    .....
      ACMWSTMGACCAACAGCACCATCTAGCGGCAGAGGAGGAA

```

FIG. 4

SIV gag -T-----A--A--A-----T-----
 SIVgagDX..-C-----G--G--C-----C-----
 #401
 AYTACCCAGTACARCARATMGGTGGTAACTAYGTCCACCT

SIV gag ----T-AAG-----AT-A--T--C-----A--AT--
 SIVgagDX..----C-GTC-----CC-G--C--T-----C--GC--
 #441
 GCCAYTRWSCCCGAGAACMYTRAAYGCTGGGTMAARYTG

SIV gag --A-----A-----A--T-----
 SIVgagDX..--C-----G-----G--C-----
 #481
 ATMGAGGARAAGAARTTYGGAGCAGAAGTAGTGCCAGGAT

SIV gag -T-----T-----T-----T-----
 SIVgagDX..-C-----C-----C-----C-----
 #521
 TYCAGGCACTGTCAGAAGGTTGCACCCCTAYGACATYAA

SIV gag T-----T-A--T--T--G-----A-----
 SIVgagDX..C-----C-G--C--C--T-----G-----
 #561
 YCAGATGYTRAAYTGYGTKGGAGACCATCARGCGGCTATG

SIV gag ----T---A-A--T--T--A-----
 SIVgagDX..----C---C-T--C--C--C-----
 #601
 CAGATYATCMGWGAYATYATMAACGAGGAGGCTGCAGATT

SIV gag -----
 SIVgagDX..-----
 #641
 GGGACTTGCAGCACCCACAACCAGCTCCACAACAAGGACA

SIV gag -----T--T-----A--T
 SIVgagDX..-----C--C-----C--C
 #681
 ACTTAGGGAGCCGTCAGGATCAGAYATYGCAGGAACMACY

SIV gag AGT-----A--T-----A-----A--A--A--
 SIVgagDX..TCC-----T--C-----G-----C--T--G--
 #721
 WSYTCAGTWGAYGAACARATCCAGTGGATGTACMGWCARC

SIV gag -----C--A-----T---A-GA-----
 SIVgagDX..-----G--C-----C---C-TC-----
 #761
 AGAACCCSATMCCAGTAGGCAACATYTACMGKMGATGGAT

SIV gag ---A-----GT----A--A--T--CA-A-----T-----A
 SIVgagDX..---G-----TC---G--G--C--TC-T-----C-----G
 #801
 CCARCTGGGKYTGCARAARTGYGYMGWATGTAYAACCCR

SIV gag --A-----
 SIVgagDX..--C-----
 #841
 ACMAACATTCTAGATGTAAAACAAGGGCCAAAAGAGCCAT

SIV gag -----
 #881
 TTCAGAGCTATGTAGACAGGTTCTACAAAAGTTTAAGAGC

SIV gag -----
 #921
 AGAACAGACAGATGCAGCAGTAAAGAATTGGATGACTCAA

SIV gag -----
 #961
 ACACTGCTGATTCAAAATGCTAACCCAGATTGCAAGCTAG

SIV gag -----
 #1001
 TGCTGAAGGGGCTGGGTGTGAATCCCACCCTAGAAGAAAT

SIV gag -----
 #1041
 GCTGACGGCTTGTC AAGGAGTAGGGGGGCGGGACAGAAG

SIV gag -----
 #1081
 GCTAGATTAATGGCAGAAGCCCTGAAAGAGGCCCTCGCAC

SIV gag -----
 #1121
 CAGTGCCAATCCCTTTTGCAGCAGCCCAACAGAGGGGACC

SIV gag -----
 #1161
 AAGAAAGCCAATTAAGTGTTGGAATTGTGGGAAGAGGGGA

SIV gag -----
#1201
CACTCTGCAAGGCAATGCAGAGCCCCAAGAAGACAGGGAT

SIV gag -----
#1241
GCTGGAAATGTGGAAAAATGGACCATGTTATGGCCAAATG

SIV gag -----
#1281
CCCAGACAGACAGGCGGGTTTTTTAGGCCTTGGTCCATGG

SIV gag -----
#1321
GGAAAGAAGCCCCGCAATTTCCCCATGGCTCAAGTGCATC

SIV gag -----
#1361
AGGGGCTGATGCCAACTGCTCCCCCAGAGGACCCAGCTGT

SIV gag -----
#1401
GGATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG

SIV gag -----
#1441
AGAGAAAAGCAGAGAGAAAGCAGAGAGAAGCCTTACAAGG

SIV gag -----
#1481
AGGTGACAGAGGATTTGCTGCACCTCAATTCTCTCTTTGG

SIV gag -----
#1521
AGGAGACCAGTAG

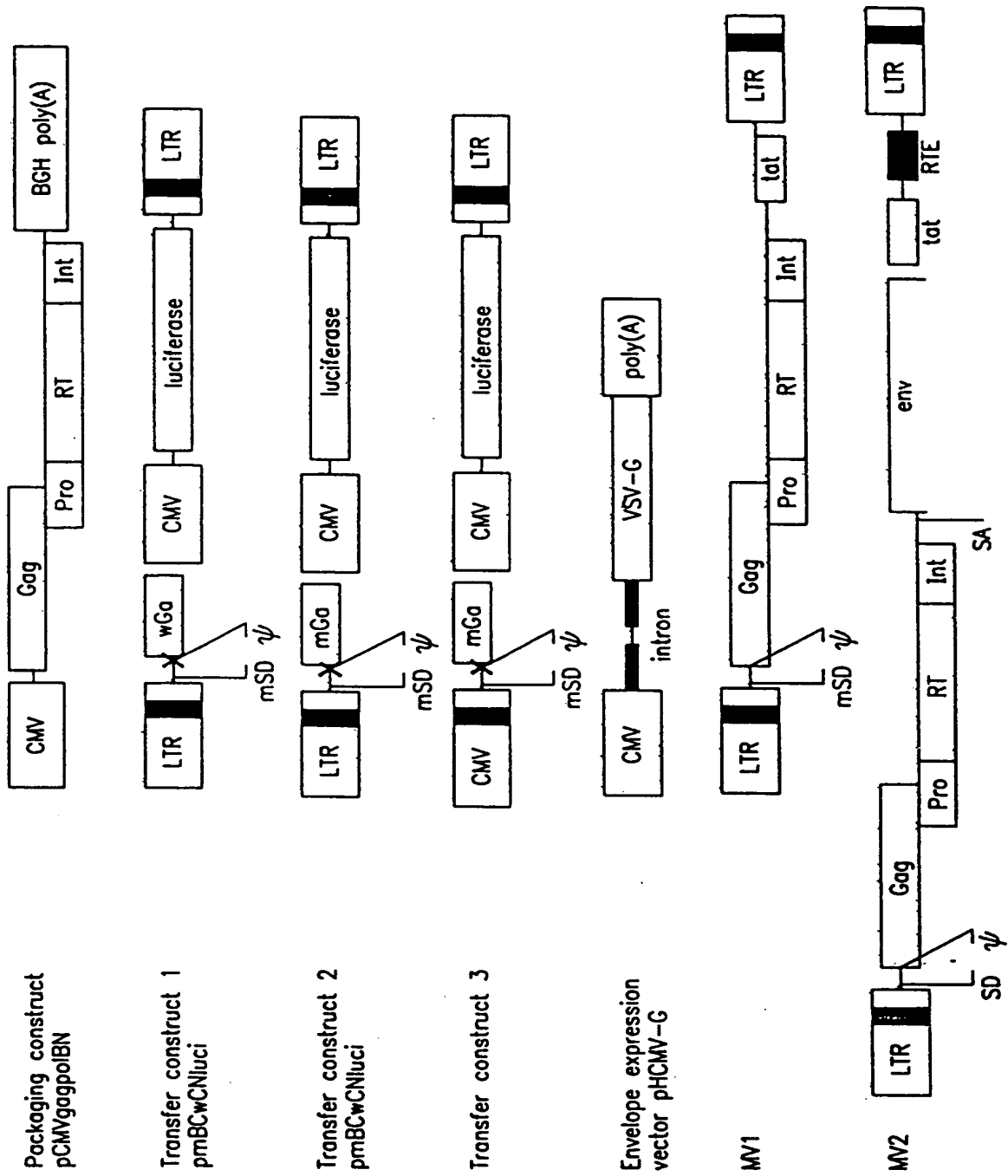


FIG. 5

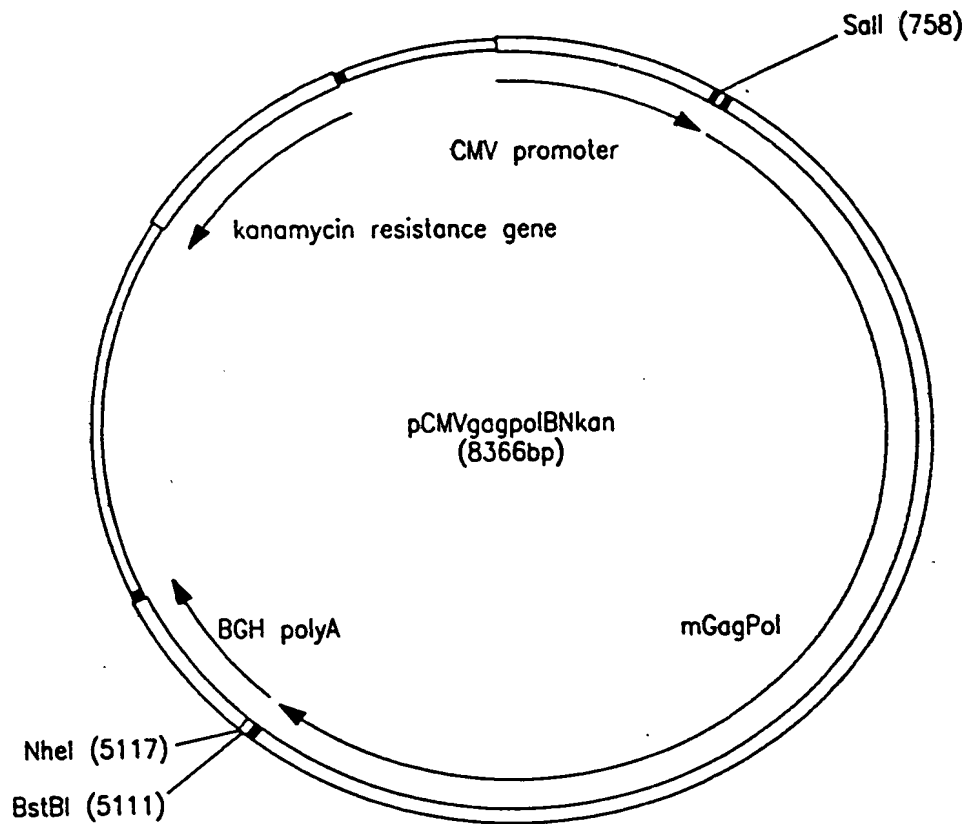


FIG. 6

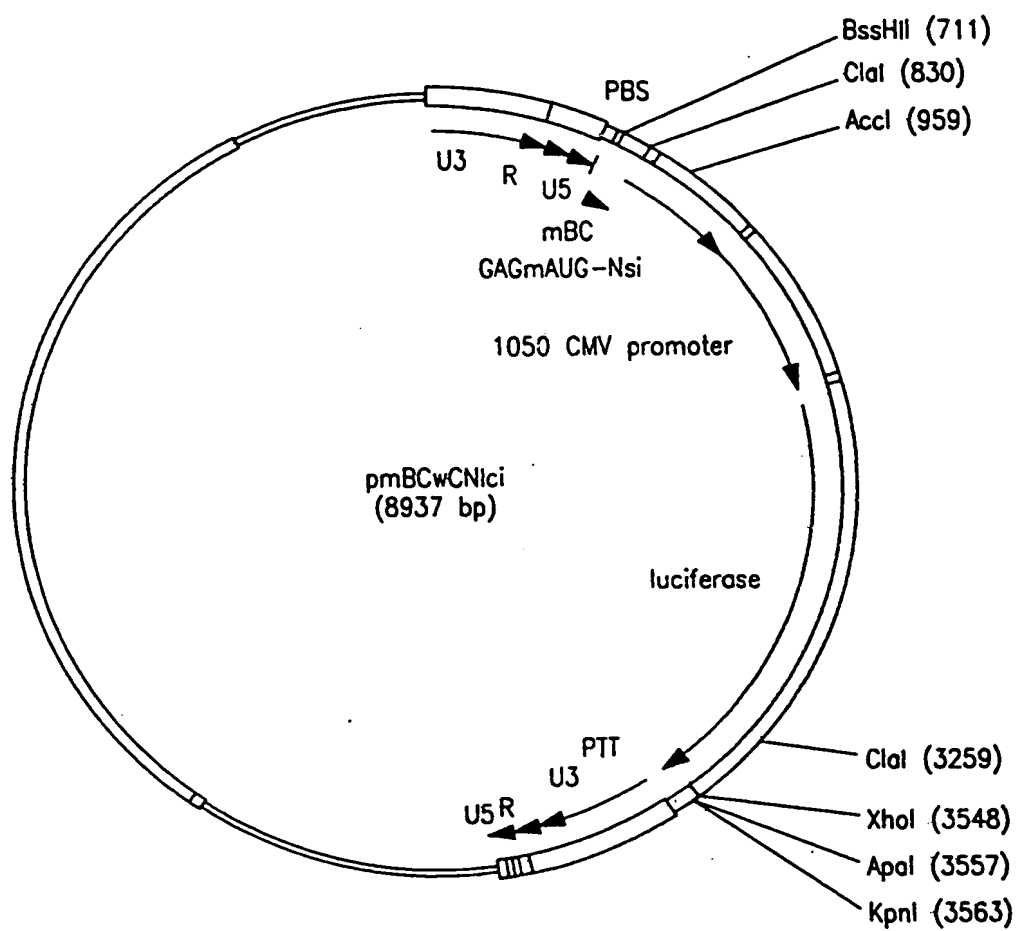


FIG. 7

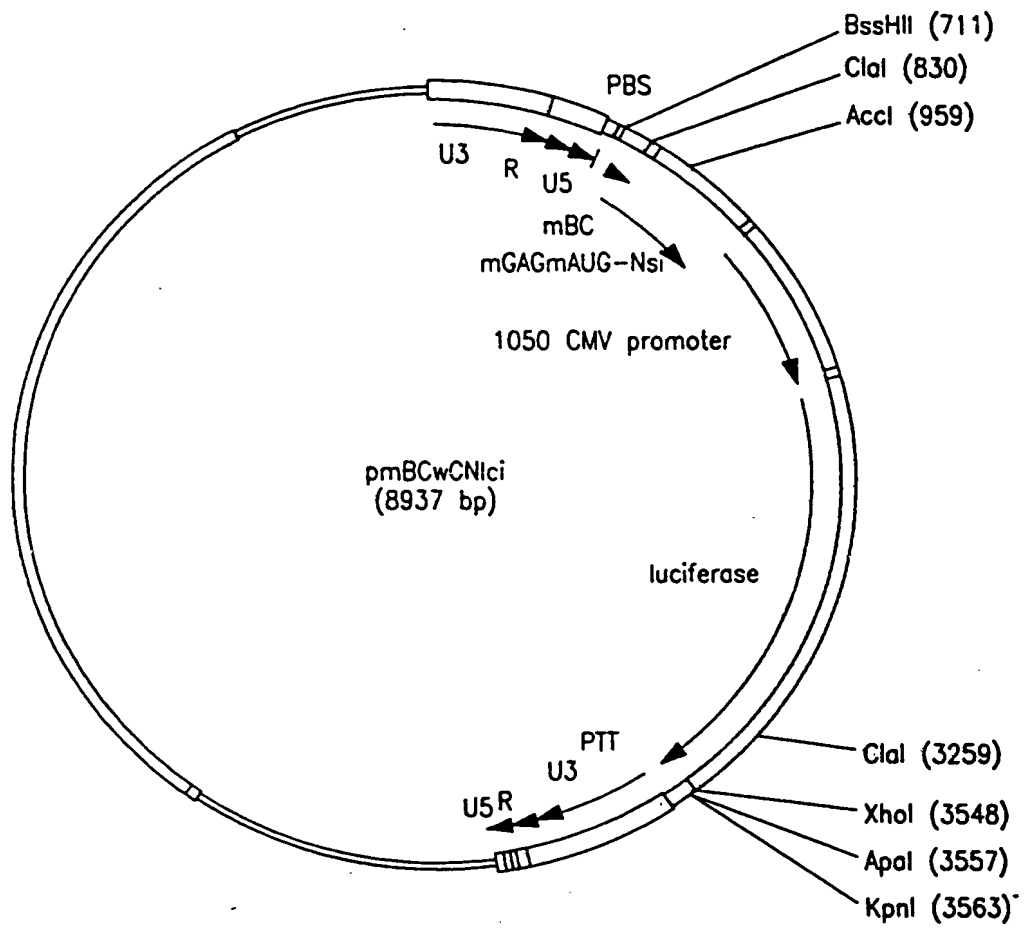


FIG. 8

```

1  CCTGGCCATT GCATACGTTG TATCCATATC ATAATATGTA CATTIATATT GGCTCATGTC CAACATTACC
71  GCCATGTTGA CATTGATTAT TGA CTAGTTA TTAATAGTAA TCAATTACGG GGTCAATTAGT TCATAGCCCA
141 TATATGGAGT TCCGCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC GACCCCGGCC
211 CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT TTCCATTGAC GTCAATGGGT
281 GGAGTATTTA CCGTAAACTG CCCACTTGGC AGTACATCAA GTGTATCATA TGCCAAGTAC GCCCCCTATT
351 GACGTCAATG ACGGTAAATG GCGCGCCTGG CATTATGCCC AGTACATGAC CTTATGGGAC TITCCTACTT
421 GGCAGTACAT CTACGTATTA GTCATCGCTA TTACCATGGT GATGCGGTTT TGGCAGTACA TCAATGGGCG
491 TGGATAGCGG TTGACTCAC GGGGATTTC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG
561 CACCAAAATC AACGGGACTT TCCAAAATGT CGTAACAACT CCGCCCCATT GACGCAAATG GGCGGTAGGC
631 GTGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG ATCGCCTGGA GACGCCATCC

                               Sall      (758)
701 ACGCTGTTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGGG GCGCGCGGTC GACAGAGAGA
771 TGGGTGCGAG AGCGTCAGTA TTAAGCGGGG GAGAATTAGA TCGATGGGAA AAAATTTCGGT TAAGGCCAGG
841 GGGAAAGAAG AAGTACAAGC TAAAGCACAT CGTATGGGCA AGCAGGGAGC TAGAACGATT CGCAGTTAAT
911 CCTGGCCTGT TAGAAACATC AGAAGGCTGT AGACAAATAC TGGGACAGCT ACAACCATCC CTTCAGACAG
981 GATCAGAGGA GCTTCGATCA CTATACAACA CAGTAGCAAC CCTCTATTGT GTGCACCAGC GGATCGAGAT
1051 CAAGGACACC AAGGAAGCTT TAGACAAGAT AGAGGAAGAG CAAAACAAGT CCAAGAAGAA GGCCAGCAG
1121 GCAGCAGCTG ACACAGGACA CAGCAATCAG GTCAGCCAAA ATTACCCTAT AGTGCAGAAC ATCCAGGGGC
1191 AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAATGC ATGGGTAAAA GTAGTAGAAG AGAAGGCTTT
1261 CAGCCCAGAA GTGATACCCA TGTTTTCAGC ATTATCAGAA GGAGCCACCC CACAGGACCT GAACACGATG
1331 TTGAACACCG TGGGGGGACA TCAAGCAGCC ATGCAAATGT TAAAGAGAC CATCAATGAG GAAGCTGCAG
1401 AATGGGATAG AGTGCATCCA GTGCATGCAG GGCCTATTGC ACCAGGCCAG ATGAGAGAAC CAAGGGGAAG
1471 TGACATAGCA GGAAC TACTA GTACCCTTCA GGAACAAATA GGATGGATGA CAAATAATCC ACCTATCCCA
1541 GTAGGAGAGA TCTACAAGAG GTGGATAATC CTGGGATTGA ACAAGATCGT GAGGATGTAT AGCCCTACCA
1611 GCATTCTGGA CATAAGACAA GGACCAAAGG AACCTTTAG AGACTATGTA GACCGGTTCT ATAAACTCT
1681 AAGAGCTGAG CAAGCTTCAC AGGAGGTAAA AAATTGGATG ACAGAAACCT TGTGGTCCA AAATGCGAAC
1751 CCAGATTGTA AGACCATCCT GAAGGCTCTC GGCCAGCGG CTACACTAGA AGAAATGATG ACAGCATGTC
1821 AGGGAGTAGG AGGACCCGGC CATAAGGCAA GAGTTTTGGC CGAGGCGATG AGCCAGGTGA CGAACTCGGC

```

FIG. 9A

1891 GACCATAATG ATGCAGAGAG GCAACTTCCG GAACCAGCGG AAGATCGTCA AGTGCTTCAA TTGTGCCAAA
 1961 GAAGGGCACA CCGCCAGGAA CTGCCGGGCC CCCCAGGAAGA AGGGCTGTTG GAAATGTGGA AAGGAAGGAC
 2031 ACCAAATGAA AGATTGTACT GAGAGACAGG CTAATTTTTT AGGGAAGATC TGGCCTTCCT ACAAGGGAAG
 2101 GCCAGGGAAT TTTCTTCAGA GCAGACCAGA GCCAACAGCC CCACCAGAAG AGAGCTTCAG GTCTGGGGTA
 2171 GAGACAACAA CTCCCCCTCA GAAGCAGGAG CCGATAGACA AGGAACTGTA TCCTTTAACT TCCCTCAGAT
 2241 CACTCTTTGG CAACGACCCC TCGTCACAGT AAGGATCGGG GGGCAACTCA AGGAAGCGCT GCTCGATACA
 2311 GGAGCAGATG ATACAGTATT AGAAGAAATG AGTTTGCCAG GAAGATGGAA ACCAAAAATG ATAGGGGGGA
 2381 TCGGGGGCTT CATCAAGGTG AGGCAGTACG ACCAGATACT CATAGAAATC TGTGGACATA AAGCTATAGG
 2451 TACAGTATTA GTAGGACCTA CACCTGTCAA CATAATTGGA AGAAATCTGT TGACCCAGAT CGGCTGCACC
 2521 TTGAACTTCC CCATCAGCCC TATTGAGACG GTGCCCCTGA AGTTGAAGCC GGGGATGGAC GGCCCCAAGG
 2591 TCAAGCAATG GCCATTGACG AAAGAGAAGA TCAAGGCCTT AGTCGAAATC TGTACAGAGA TGGAGAAGGA
 2661 AGGGAAGATC AGCAAGATCG GGCCTGAGAA CCCCTACAAC ACTCCAGTCT TCGCAATCAA GAAGAAGGAC
 2731 AGTACCAAGT GGAGAAAGCT GGTGGACTTC AGAGAGCTGA ACAAGAGAAC TCAGGACTTC TGGGAAGTTC
 2801 AGCTGGGCAT CCCACATCCC GCTGGGTTGA AGAAGAAGAA GTCAGTGACA GTGCTGGATG TGGGTGATGC
 2871 CTAATTCTCC GTTCCCTTGG ACGAGGACTT CAGGAAGTAC ACTGCCTTCA CGATACCTAG CATCAACAAC
 2941 GAGACACCAG GCATCCGCTA CCAGTACAAC GTGCTGCCAC AGGGATGGAA GGCATACCA GCCATCTTTC
 3011 AAAGCAGCAT GACCAAGATC CTGGAGCCCT TCCGCAAGCA AAACCCAGAC ATCGTGATCT ATCAGTACAT
 3081 GGACGACCTC TACGTAGGAA GTGACCTGGA GATCGGGCAG CACAGGACCA AGATCGAGGA GCTGAGACAG
 3151 CATCTGTGA GGTGGGGACT GACCACACCA GACAAGAAGC ACCAGAAGGA ACCTCCCTTC CTGTGGATGG
 3221 GCTACGAACT GCATCCTGAC AAGTGACAG TGCAGCCCAT CGTGCTGCCT GAGAAGGACA GCTGGACTGT
 3291 GAACGACATA CAGAAGCTCG TGGGCAAGTT GAACTGGGCA AGCCAGATCT ACCCAGGCAT CAAAGTTAGG
 3361 CAGCTGTGCA AGCTGCTTCG AGGAACCAAG GCACTGACAG AAGTGATCCC ACTGACAGAG GAAGCAGAGC
 3431 TAGAACTGGC AGAGAACCGA GAGATCCTGA AGGAGCCAGT ACATGGAGTG TACTACGACC CAAGCAAGGA
 3501 CCTGATCGCA GAGATCCAGA AGCAGGGGCA AGGCCAATGG ACCTACCAAA TCTACCAGGA GCCCTTCAAG
 3571 AACCTGAAGA CAGGCAAGTA CGCAAGGATG AGGGGTGCCC ACACCAACGA TGTGAAGCAG CTGACAGAGG
 3641 CAGTGCAGAA GATCACCACA GAGAGCATCG TGATCTGGGG CAAGACTCCC AAGTTCAAGC TGCCCATACA
 3711 GAAGGAGACA TGGGAGACAT GGTGGACCGA GTA CTGGCAA GCCACCTGGA TCCCTGAGTG GGAGTTCGTG

FIG. 9B

3781 AACACCCCTC CTTGGTGAA ACTGTGGTAT CAGCTGGAGA AGGAACCCAT CGTGGGAGCA GAGACCTTCT
 3851 ACGTGGATGG GGCAGCCAAC AGGAGACCA AGCTGGGCAA GGCAGGCTAC GTGACCAACC GAGGACGACA
 3921 GAAAGTGGTG ACCCTGACTG ACACCACCAA CCAGAAGACT GAGCTGCAAG CCATCTACCT AGCTCTGCAA
 3991 GACAGCGGAC TGGAAGTGAA CATCGTGACA GACTCACAGT ACGCACTGGG CATCATCCAA GCACAACCAG
 4061 ACCAATCCGA GTCAGAGCTG GTGAACCAGA TCATCGAGCA GCTGATCAAG AAGGAGAAAG TGTACCTGGC
 4131 ATGGGTACCA GCACACAAAG GAATTGGAGG AAATGAACAA GTAGATAAAT TAGTCAGTGC TGGGATCCGG
 4201 AAGGTGCTGT TCCTGGACGG GATCGATAAG GCCCAAGATG AACATGAGAA GTACCACTCC AACTGGCGCG
 4271 CTATGGCCAG CGACTTCAAC CTGCCACCTG TAGTAGCAAA AGAAATAGTA GCCAGCTGTG ATAAATGTCA
 4341 GCTAAAAGGA GAAGCCATGC ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAGCTGGA CTGCACGCAC
 4411 CTGGAGGGGA AGGTGATCCT GGTAGCAGTT CATGTAGCCA GTGGATATAT AGAAGCAGAA GTTATCCCTG
 4481 CTGAAACTGG GCAGGAAACA GCATATTTTC TTTTAAATT AGCAGGAAGA TGGCCAGTAA AAACAATACA
 4551 CACGGACAAC GGAAGCAACT TCACTGGTGC TACGGTTAAG GCCGCCTGTT GGTGGGCGGG AATCAAGCAG
 4621 GAATTTGGAA TTCCCTACAA TCCCAATCG CAAGGAGTCG TGGAGAGCAT GAACAAGGAG CTGAAGAAGA
 4691 TCATCGGACA AGTGAGGGAT CAGGCTGAGC ACCTGAAGAC AGCAGTGCAG ATGGCAGTGT TCATCCACAA
 4761 CTTCAAAAGA AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG GAAAGGATCG TGGACATCAT CGCCACCGAC
 4831 ATCCAAACCA AGGAGCTGCA GAAGCAGATC ACCAAGATCC AGAACTTCCG GGTGTACTAC CGCGACAGCC
 4901 GCAACCCACT GTGGAAGGGA CCAGCAAAGC TCCTCTGGAA GGGAGAGGGG GCAGTGGTGA TCCAGGACAA
 4971 CAGTGACATC AAAGTGGTGC CAAGGCGCAA GGCCAAGATC ATCCGCGACT ATGGAAAACA GATGCCAGGT
 5041 GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTAGAACC TGAAGAGCC TGGTGAAGCA CCATATGGCG


 NheI (5117)
 BstBI (5111)
 5111 TTCGAAGCTA GCCTCGAGAT CCAGATCTGC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT TTGCCCCTCC
 5181 CCCGTGCCTT CTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT
 5251 CGCATTGTCT GAGTAGGTGT CATTCTATTG TGGGGGGTGG GGTGGGGCAG CACAGCAAGG GGGAGGATTG
 5321 GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT ATGCGTACCC AGGTGCTGAA GAATTGACCC
 5391 GGTTCCTCCT GGGCCAGAAA GAAGCAGGCA CATCCCCTTC TCTCTGACAC ACCCTGTCCA CGCCCCTGGT
 5461 TCTTAGTTCC AGCCCCACTC ATAGGACACT CATAGCTCAG GAGGGCTCCG CCTTCAATCC CACCCGCTAA
 5531 AGTACTTGA GCGGTCTCTC CCTCCCTCAT CAGCCCACCA AACCAAACCT AGCCTCCAAG AGTGGGAAGA

FIG. 9C

5601 AATTAAAGCA AGATAGGCTA TTAAGTGCAG AGGGAGAGAA AATCCCTCCA ACATGTGAGG AAGTAATGAG

5671 AGAAATCATATA GAATTTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTTCGTTCCG CTGCGGCGAG

5741 CCGGTATCAGC TCACTCAAAG GCGGTAATAC GGTATATCCAC AGAATCAGGG GATAACGCAG GAAAGAACAT

5811 GTGAGCAAAA GGGCAGCAAA AGGCCAGGAA CCGTAAAAAG GCCCGCTTGC TGGCGTTTTT CCATAGGCTC

5881 CGCCCCCTG ACAGGCATCA CAAAAATCGA CGCTCAAGTC ACAGGTGGCG AAACCCGACA GGAATAATAA

5951 GATACCAGGC GTTTCCCTCT GGAAGCTCCC TCCTGCTCCG ACCCTGCCGC TTACCGGATA

6021 CCTGTCCGCC TTTCTCCCTT CCGGAAGCGT GCGGCTTTCT CAATGCTCAC GCTGTAGGTA TCTCAGTTCC

6091 GTGTAGGTCG TTCGCTCCAA GCTCGGCTGT GTGCACGAAC CCCCCGTTC A GCGCGACCGC TCGCCCTTAT

6161 CCGGTAACTA TCCTCTTGAG TCCAACCCGG CTTATCGCCA CTGGCAGCAG CCACTGGTAA

6231 CAGGATTAGC AGAGCGAGGT ATGTAGCGCG TCTACACAG TTCTTGAAGT GGTGGCCTAA CTACCGGTAC

6301 ACTAGAAGCA CAGTATTGG TATCTGCGCT CTGCTGAAGC CAGTTACCTT CCGAAAAAGA GTTGGTAGCT

6371 CTTGATCCCG CAAACAAACC ACCGCTGGTA TGTATGAGT TTTTGTTCG AAGCAGCAGG TTACCGCGCAG

6441 AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGTCTGACG CTCAGTGGAA CGAAAACTCA

6511 CGTTAAGGGA TTTTGGTCAT CAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT TAAAAATGAA

6581 GTTTTAAATC AATCTAAAGT ATATATGAGT AATCTTGCTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC

6651 ACCTATCTCA GCGATCTGTC TATTTCTGTT ATCCATAGTT GCCTGACTCC GGGGGGGGGG GCGCGTGAGG

6721 TCTGCCTCGT GAAGAAGGTG TTGCTGACTC ATACCAGGCC TGAATCGCCC CATCATCCAG CCAGAAAGTG

6791 AGGGAGCCAC GGTGTATGAG AGCTTTGTTG TAGGTGGACC ACTTGGTGAT TTTGAACCTT TGCTTTGCCA

6861 CGGAACGGTC TCGCTTGTCG GGAAGATGCG CTTCAACTCA AGTGTTACAA GCAAAAGTTC GATTATTATCA

6931 ACAAAGCCGC CGTCCCGTCA AGTCAGCGTA ATGCTCTGCC GCAATTTATT CATATCAGGA TTATCAATAC CATATTTTGT

7001 AGAAAACTC ATCGAGCATC AAATGAAACT euHisPheG1 MetAspProA snAspIleG1 yTyrLysGln

271 PhePheGlu AspLeuMetL TTCTGTAATG AAGGAGAAAA CTCACCGAGG CAGTTCCATA GGATGGCAAG ATCTCGGTAT

7071 AAAAAGCCGT ysGlnLeuSe rProSerPhe GluGlyLeuC ysAsnTrpLe uIleAlaLeu AspGlnTyrA

248 PheLeuArgL TTCCGACTCG TCCAACATCA ATACAACCTA TTAATTTCCC CTGCTCAAAA ATAAGGTTAT

7141 CCGTCTGCGA eGlyValArg GlyValAspI leCysGlyI1 eLeuLysGly GluAspPheI leLeuAsnAs

224 rgAspAlaI1 ATCACCATGA GTGACGACTG AATCCGGTGA GAATGGCAAA AGCTTATGCA TTTCTTTCCA

7211 CAAGTGAGAA ATCACCATGA hrValValSe rAspProSer PheProLeuL euLysHisMe tGluLysTrp

201 pLeuSerPhe AspGlyHisT CATTACGCTC yAsnArgGlu AspAspPheA spSerAlaAs pValLeuGly AsnAsnMetA

7281 GACTTGTTC A CAGGCCAGC CATTACGCTC yAsnArgGlu AspAspPheA spSerAlaAs pValLeuGly AsnAsnMetA

178 ValGlnGluV alProTrpG1 ACGAAATACG CGATCGCTGT TAAAGGACA ATTACAAACA GGAATCGAAT

7351 CGTGATTGCG CCTGACCGAG ArgPheValA rgAspSerAs nPheProCys AsnCysValP roIleSerHi

154 rgSerGlnAl aGlnAlaLeu GCCACGCGAT CAACAATATT TTCACCTGAA TCAGGATATT CTCTAATAC

7421 GCAACCGCGG LeuPheValA laLeuAlaAs pValIleAsn GluGlySera spProTyrGl uGluLeuVal

131 sLeuArgArg LeuPheValA laLeuAlaAs pValIleAsn GluGlySera spProTyrGl uGluLeuVal

7491 CTGGAATGCT GTTTTCCCGG CGATCGCAGT GGTGAGTAAC CATGCATCAT CAGGAGTACG GATAAAATGC

108 GlnPheAlaT hrLysGlyPr oIleAlaThr ThrLeuLeuT rpAlaAspAs pProThrArg ilePheHisL

7561 TTGATGGTGC GAAGAGGCAT AAATTCCGTC AGCCAGTTTA GTCTGACCAT CTCATCTGTA ACATCATTTGG

84 ysIleThrPr oLeuProMet PheGluThrL euTrpAsnLe uArgValMet GluAspThrV alAspAsnAl

7631 CAACGCTACC TTTGCCATGT TTCAGAAACA ACTCTGGCGC ATCGGGCTTC CCATACAATC GATAGATTGT

61 aValSerGly LysGlyHisL ysLeuPheLe uGluProAla AspProLysG lyTyrLeuAr gTyrIleThr

7701 CGCACCTGAT TGCCCCACAT TATCGCGAGC CCATTTATAC CCATATAAAT CAGCATCCAT GTTGGAAATT

38 AlaGlySerG lnGlyValAs nAspArgAla TrpLysTyrG lyTyrLeuAs pAlaAspMet AsnSerAsnL

7771 AATCGCGGCC TCGAGCAAGA CGTTTCCCGT TGAATATGGC TCATAACACC CTTGTATTA CTGTTTATGT

14 euArgProAr gSerCysSer ThrGluArgG CATGATGATA TATTTTATC TTGTGCAATG TAACATCAGA GATTTTGAGA

7841 AAGCAGACAG TTTTATTGTT CCCCCCATT TTTTATTTATC TATCAGGTTT ATTGTCTCAT GAGCGGATAC

7911 CACAACGTGG CTITCCCCC CCCCCCATT TTTTATTTATC TATCAGGTTT ATTGTCTCAT GAGCGGATAC

7981 ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT CCGGCACATT TCCCCGAAAA GTGCCACCTG

8051 ACGTCTAAGA AACCATTTAT ATCATGACAT TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGTCT

8121 CGCGCGTTTC GGTGATGACG GTGAAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG

8191 TAAGCGGATG CCGGAGCAGC ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC

8261 TTAACATATGC GGCATCAGC CAGATTGTAC GTAGATGCA CCATATGCGG TGTGAAATAC CGCACAGATG

8331 CGTAAGGAGA AAATACCGCA TCAGATTGGC TATTGG (SEQUENCE ID NO: 6)

FIG. 9D

1 TGGAAGGGCT AATTTGGTCC CAAAAAAGAC AAGAGATCCT TGATCTGTGG ATCTACCACA CACAAGGCTA
 71 CTTCCCTGAT TGGCAGAACT ACACACCAGG GCCAGGGATC AGATATCCAC TGACCTTTGG ATGGTGCTTC
 141 AAGTTAGTAC CAGTTGAACC AGAGCAAGTA GAAGAGGCCA AATAAGGAGA GAAGAACAGC TTGTTACACC
 211 CTATGAGCCA GCATGGGATG GAGGACCCGG AGGGAGAAGT ATTAGTGTGG AAGTTTGACA GCCTCCTAGC
 281 ATTTCTGCAC ATGGCCCGAG AGCTGCATCC GGAGTACTAC AAAGACTGCT GACATCGAGC TTTCTACAAG
 351 GGACTTTCCG CTGGGGACTT TCCAGGGAGG TGTGGCCTGG GCGGCACTGG GGAGTGGCGA GCCCTCAGAT
 421 GCTACATATA AGCAGCTGCT TTTTGCCTGT ACTGGGTCTC TCTGGTTAGA CCAGATCTGA GCCTGGGAGC
 491 TCTCTGGCTA ACTAGGGAAC CCACTGCTTA AGCCTCAATA AAGCTTGCCT TGAGTGCTCA AAGTAGTGTC
 561 TGCCCGTCTG TTGTGTGACT CTGGTAACTA GAGATCCCTC AGACCCTTTT AGTCAGTGTG GAAAATCTCT
 631 AGCAGTGGCG CCCGAACAGG GACTTGAAAG CGAAAGTAAA GCCAGAGGAG ATCTCTCGAC GCAGGACTCG
 701 GCTTGCTGAA ^{BssHII (711)} GCGCGCacgg caagaggcga ggggcggcgC ctgACgagGa cgccaaaaat tttgactagc
 771 ggaggctaga aggagagagC ^{ClaI (830)} TCGGTGCGAG AGCGTCAGTA TCAAGCGGGG GAGAATTAGA TCGATGGGAA
 841 AAAATTTCGGT TAAGGCCAGG GGGAAAGAAA AAATATAAAT TAAAACATAT AGTATGGGCA AGCAGGGAGC
 911 TAGAACGATT CGCAGTTAAT CCTGGCCTGT TAGAAACATC AGAAGGCTGT AGACAAATAC TGGGACAGCT
 981 ACAACCATCC CTTAGACAG GATCAGAAGA ACTTAGATCA TTATATAATA CAGTAGCAAC CCTCTATTGT
 1051 GTGCATCAAA GGATAGAGAT AAAAGACACC AAGGAAGCTT TAGACAAGAT AGAGGAAGAG CAAAACAAAA

FIG. 10A

1121 GTAAGAAAAA AGCACAGCAA GCACGAGCTG ACACAGGACA CAGCAATCAG GTCAGCCAAA ATTACCCTAT
 1191 AGTGCAGAAC ATCCAGGGGC AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAACGA TAAGCTTGGG
 1261 AGTTCCGCGT TACATAACTT ACGGTAAATG GCGCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC
 1331 GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGAGTAT
 1401 TTACGGTAAA CTGCCCCACTT GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA
 1471 ATGACGGTAA ATGCCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTGGCAGTA
 1541 CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG
 1611 CGGTTTGA CT CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG GAGTTTGTTT TGGCACCAAA
 1681 ATCAACGGGA CTTTCACAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGGCGGTA GCGGTGTACG
 1751 GTGGGAGGTC TATATAAGCA GAGCTCGTTT AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT
 1821 TTTGACCTCC ATAGAAGACA CCGACTCTAG AGgatccATC TAAGTAAGCT TGGCATTCCG GACTGTGTTG
 1891 TAAAATGGAA GACGCCAAAA ACATAAAGAA AGGCCCGGCG CCATTCTATC CTCTAGAGGA TGAACCGCT
 1961 GGAGAGCAAC TGCATAAGGC TATGAAGAGA TACGCCCTGG TTCCTGGAAC AATTGCTTTT ACAGATGCAC
 2031 ATATCGAGGT GAACATCAGC TACGCGGAAT ACTTCGAAAT GTCCGTTCCG TTGGCAGAAG CTATGAAACG
 2101 ATATGGGCTG AATACAAATC ACAGAATCGT CGTATGCAGT GAAAACCTCTC TTCAATTCTT TATGCCGGTG
 2171 TTGGGCCCCG TATTTATCGG AGTTGCAGTT GCGCCCGCGA ACGACATTTA TAATGAACGT GAATTGCTCA
 2241 ACAGTATGAA CATTTCGCAG CCTACCGTAG TGTGTTGTTT CAAAAAGGGG TTGCAAAAAA TTTTGAACGT
 2311 GCAAAAAAAA TTACCAATAA TCCAGAAAAT TATTATCATG GATTCTAAAA CGGATTACCA GGGATTTCAG

FIG. 10B

2381 TCGATGTACA CGTTCGTAC ATCTCATCTA CCTCCCGGTT TTAATGAATA CGATTTTGTA CCAGAGTCCT
 2451 TTGATCGTGA CAAAACAATT GCACTGATAA TGAATTCCTC TGGATCTACT GGGTACCTA AGGGTGTGGC
 2521 CCTTCCGCAT AGAACTGCCT GCGTCAGATT CTCGCATGCC AGAGATCCTA TTTTGGCAA TCAAATCATT
 2591 CCGGATACTG CGATTTTAAG TGTGTGTTCCA TTCCATCAGG GTTTTGAAT GTTTACTACA CTCGGATATT
 2661 TGATATGTGG ATTTGAGTGC GTCTTAATGT ATAGATTIGA AGAAGAGCTG TTTTACGAT CCCTTCAGGA
 2731 TTACAAAATT CAAAGTGCGT TGCTAGTACC AACCTATTT TCATTCTTCG CAAAAGCAC TCTGATTGAC
 2801 AAATACGATT TATCTAATTT ACACGAAATT GCTTCTGGGG GCGCACCTCT TTCGAAAGAA GTCCGGCAAG
 2871 CGGTTGCAAA ACGCTTCCAT CTTCCAGGGA TACGACAAGG ATATGGGCTC ACTGAGACTA CATCAGCTAT
 2941 TCTGATTACA CCCGAGGGGG ATGATAAACC GGGCGCGGTC GGTAAAGTTG TTCCATTTTT TGAAGCGAAG
 3011 GTTGTGGATC TGGATACCGG GAAAACGCTG GCGGTTAATC AGAGAGGCGA ATTATGTGTC AGAGGACCTA
 3081 TGATTATGTC CGGTTATGTA AACAATCCGG AAGCGACCAA CGCCTTGATT GACAAGGATG GATGGCTACA
 3151 TTCTGGAGAC ATAGCTTACT GGGACGAAGA CGAACACTTC TTCATAGTTG ACCGCTTGAA GTCTTTAATT
 3221 AAATACAAAG GATATCAGGT GGGCCCCGCT GAATTGGAAT CGATATTGTT ACAACACCCC AACATCTTCG
 3291 ACCCGGGCGT GGCAGGTCTT CCCGACGATG ACGCCGGTGA ACTTCCCGCC GCCGTTGTTG TTTGGAGCA
 3361 CGGAAAGACG ATGACGGAAG AAGAGATCGT GGATTACGTC GCCAGTCAAG TAACAACCGC GAAAAAGTTG
 3431 CGCGGAGGAG TTGTGTTTGT GGACGAAGTA CCGAAAGGTC TTACCGGAAA ACTCGACGCA AGAAAAATCA
 3501 GAGAGATCCT CATAAAGGCC AAGAAGGCCG GAAAGTCCAA ATTGTAACTC GAGGGGGGGC CCGGTACCTT
 XhoI (3548) ApaI (3557)
 KpnI (3563)

FIG. 10C

3571 TAAGACCAAT GACTTACAAG GCAGCTGTAG ATCTTAGCCA CTTTTTAAAA GAAAAGGGGG GACTGGAAGG
 3641 GCTAATTAC TCCCAAAGAA GACAAGATAT CCTTGATCTG TGGATCTACC ACACACAAGG CTACTTCCCT
 3711 GATTGGCAGA ACTACACACC AGGGCCAGGG GTCAGATATC CACTGACCTT TGGATGGTGC TACAAGCTAG
 3781 TACCAGTTGA GCCAGATAAG GTAGAAGAGG CCAATAAAGG AGAGAACACC AGCTTGTTAC ACCCTGTGAG
 3851 CCTGCATGGA ATGGATGACC CTGAGAGAGA AGTGTTAGAG TGGAGGTTTG ACAGCCGCCT AGCATTTCAT
 3921 CACGTGGCCC GAGAGCTGCA TCCGGAGTAC TTCAAGAACT GCTGACATCG AGCTTGCTAC AAGGGACTTT
 3991 CCGCTGGGGA CTTTCCAGGG AGGCGTGGCC TGGGCGGGAC TGGGAGTGG CGAGCCCTCA GATGCTGCAT
 4061 ATAAGCAGCT GCTTTTGGC TGTACTGGGT CTCTCTGGTT AGACCAGATC TGAGCCTGGG AGCTCTCTGG
 4131 CTAAGTAGGG AACCCACTGC TTAAGCCTCA ATAAAGCTTG CCTTGAGTGC TTCAAGTAGT GTGTGCCCCG
 4201 CTGTTGTGTG ACTCTGGTAA CTAGAGATCC CTCAGACCCT TTAGTCACT GTGGAAAATC TCTAGCACCC
 4271 CCCAGGAGGT AGAGGTTGCA GTGAGCCAAG ATCGCGCCAC TGCATTCCAG CCTGGGCAAG AAAACAAGAC
 4341 TGTCTAAAAA AATAATAATA AGTTAAGGGT ATTAATATATA TTTATACATG GAGGTCATAA AAATATATAT
 4411 ATTTGGGCTG GCGCGAGTGG CTCACACCTG CGCCCGGGCC TTTGGGAGGC CGAGGCAGGT GGATCACCTG
 4481 AGTTTGGGAG TTCCAGACCA GCCTGACCAA CATGGAGAAA CCCCTTCTCT GTGTATTTT ATGAGATTT
 4551 ATTTTATGTG TATTTTATTC ACAGGTATTT CTGGAAGAACT GAAACTGTTT TTCTCTACT CTGATACCAC
 4621 AAGAATCATC AGCACAGAGG AAGACTTCTG TGATCAAATG TGGTGGGAGA GGGAGGTTTT CACCAGCACA
 4691 TGAGCAGTCA GTTCTGCCGC AGACTCGGCG GGTGTCCTTC GGTTCACTTC CAACACCGCC TGCCTGGAGA
 4761 GAGGTCAGAC CACAGGGTGA GGGCTCAGTC CCCAAGACAT AAACACCCAA GACATAAACA CCCAACAGGT
 4831 CCACCCCGCC TGCTGCCAG GCAGAGCCGA TTCACCAAGA CGGGAATTAG GATAGAGAAA GAGTAACTCA
 4901 CACAGAGCCG GCTGTGCGGG AGAACGGAGT TCTATTATGA CTCAAATCAG TCTCCCCAAG CATTCCGGGA
 4971 TCAGAGTTTT TAAGGATAAC TTAGTGTGTA GGGGGCCAGT GAGTTGGAGA TGAAGCGTA GGGAGTCGAA
 5041 GGTGTCTTTT TCGCGCCAGT CAGTTCCTGG GTGGGGGCCA CAAGATCGGA TGAGCCAGTT TATCAATCCG
 5111 GGGGTGCCAG CTGATCCATG GAGTGCAGGG TCTGCAAAAT ATCTCAAGCA CTGATTGATC TTAGGTTTTA
 5181 CAATAGTGAT GTTACCCAG GAACAATTG GGAAGGTCA GAATCTTGTA GCCTGTAGCT GCATGACTCC
 5251 TAAACCATAA TTTCTTTTTT GTTTTTTTTT TTTTATTTTT GAGACAGGGT CTCACTCTGT CACCTAGGCT
 5321 GGAGTGCAGT GGTGCAATCA CAGTCACTG CAGCCCCTAG AGCGCGCGCC ACCCGGTGG AGCTCCAATT
 5391 CGCCCTATAG TGAGTCGTAT TACAATTCAC TGGCCGTCGT TTTACAACGT CGTGAAGTGG AAAACCCCTGG
 5461 CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC
 5531 ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCGA AATTGTAAAC GTTAATATT
 5601 TGTTAAAAAT CGCGTTAAAT TTTTGTAAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT
 5671 CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGTTCCAG TTTGGAACAA GAGTCCACTA
 5741 TTAAAGAACG TGGACTCCAA CGTCAAAGCG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC
 5811 CATCACCTTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA AGCACTAAAT CGGAACCTTA AAGGGAGCCC
 5881 CCGATTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG

FIG. 10D

```

5951 GCGCGTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC ACCCGCCGCG CTTAATGCGC
6021 CGCTACAGGG CGCGTCCCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTC TTTATTTTTC
6091 TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
6161 GGAACAGTAT GAGTATTCAA CATTTCGGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT
6231 TTTTGCTCAC CCAGAAACGC TGGTGAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC
6301 ATCGAACTGG ATCTCAACAG CCGTAAGATC CTGAGAGTT TTCGCCCCGA AGAAGCTTTT CCAATGATGA
6371 GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCC
6441 CCGCATACAC TATTCTCAGA ATGACTTGCT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
6511 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC TCGCGCCAAC TTACTTCTGA
6581 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCCTGA
6651 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG
6721 GCAACAACGT TGCGCAAACT ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT
6791 GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
6861 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GCGCAGATGG TAAGCCCTCC
6931 CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA
7001 TAGGTGCTC ACTGATTAA CATTGGTAAC TGTGAGACCA AGTTTACTCA TATATACTTT AGATTGATTT
7071 AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ACTTCATGAC CAAAATCCCT
7141 TAACGTGAT TTTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
7211 TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
7281 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATAACAAA TACTGTCTTT
7351 CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA
7421 TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT
7491 ACCGGATAAG GCGCAGCGGT CCGGCTGAAC GGGGGGTTCT TGCACACAGC CCAGCTTGA GCGAACGACC
7561 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG
7631 ACAGGTATCC GGTAAAGCGC AGGGTCGGA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCTG
7701 GTATCTTTAT AGTCCTGTCT GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG
7771 GGGCGGAGCC TATGGAAAAA CGCCAGCAAC CGGGCCTTTT TACGGTTTCT GGCCTTTTGC TGGCCTTTTG
7841 CTCACATGTT CTTTCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTGG AGTGAGCTGA
7911 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCCAATA
7981 CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA CGACAGGTTT CCCGACTGGA
8051 AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTAGCT CACTCATTAG GCACCCCAGG CTTTACACTT
8121 TATGCTTCCG GCTCGTATGT TGTGTGGAAT TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC
8191 CATGATTACG CCAAGCTCGG AATTAACCCT CACTAAAGGG AACAAAAGCT GCTGCAGGGT CCCTAACTGC
8261 CAAGCCCCAC AGTGTGCCCT GAGGCTGCCC CTTCCTTCTA GCGGCTGCCC CCACTCGGCT TTGCTTTCCC
8331 TAGTTTCAGT TACTTGGCTT CAGCCAAGGT CTGAAACTAG GTGCGCACAG AGCGGTAAGA CTGCGAGAGA
8401 AAGAGACCAG CTTTACAGGG GGTTTATCAC AGTGCACCCT GACAGTCGTC AGCCTCACAG GGGGTTTATC
8471 ACATTGCACC CTGACAGTCG TCAGCCTCAC AGGGGGTTTA TCACAGTGCA CCCTTACAAT CATTCCATTT
8541 GATTACAAT TTTTTTAGTC TCTACTGTGC CTAAGTTGTA AGTTAAATTT GATCAGAGGT GTGTTCCAG
8611 AGGGGAAAAAC AGTATATACA GGGTTTCAGTA CTATCGCATT TCAGGCCTCC ACCTGGGTCT TGGAAATGTT
8681 CCCCCGAGGG GTGATGACTA CCTCAGTTGC ATCTCCACAG GTCACAGTGA CACAAGATAA CCAAGACACC
8751 TCCCAAGGCT ACCACAATGG GCCGCCCTCC ACGTGCACAT GGCCGGAGGA ACTGCCATGT CGGAGGTGCA
8821 AGCACACCTG CGCATCAGAG TCCTTGCTGT GGAGGGAGGG ACCAGCGCAG CTTCCAGCCA TCCACCTGAT
8891 GAACAGAACC TAGGGAAAGC CCCAGTTCTA CTTACACCAG GAAAGGC (SEQUENCE ID NO: 8)

```

FIG. 10E

1 TGGAAGGGCT AATTGGTCC CAAAAAAGAC AAGAGATCCT TGATCTGTGG ATCTACCACA CACAAGGCTA
 71 CTTCCCTGAT TGGCAGAACT ACACACCAGG GCCAGGGATC AGATATCCAC TGACCTTTGG ATGGTGCTTC
 141 AAGTTAGTAC CAGTTGAACC AGAGCAAGTA GAAGAGGCCA AATAAGGAGA GAAGAACAGC TTGTTACACC
 211 CTATGAGCCA GCATGGGATG GAGGACCCGG AGGGAGAAGT ATTAGTGTGG AAGTTTGACA GCCTCCTAGC
 281 ATTCGTCAC ATGGCCCGAG AGCTGCATCC GGAGTACTAC AAAGACTGCT GACATCGAGC TTTCTACAAG
 351 GGACTTTCGG CTGGGGACTT TCCAGGGAGG TGTGGCCTGG GCGGGACTGG GGAGTGGCGA GCCCTCAGAT
 421 GCTACATATA AGCAGCTGCT TTTTGCTGT ACTGGGTCTC TCTGGTTAGA CCAGATCTGA GCCTGGGAGC
 491 TCTCTGGCTA ACTAGGGAAC CCACTGCTTA AGCCTCAATA AAGCTTGCCT TGAGTGCTCA AAGTAGTGTG
 561 TGCCCGTCTG TTGTGTGACT CTGGTAACTA GAGATCCCTC AGACCCTTTT ACTCAGTGTG GAAAATCTCT
 631 AGCAGTGGCG CCCGAACAGG GACTTGAAAG CGAAAGTAAA GCCAGAGGAG ATCTCTCGAC GCAGGACTCG
 701 GCTTGCTGAA ^{BssHII (711)} GCGCGCacgg caagaggcga ggggcggcgC ctgACgagGa cgccaaaaat tttgactagc
 771 ggaggctaga aggagagagC TCGGTCCGAG AGCGTCAGTA TTAAGCGGGG GAGAATTAGA TCGATGGGAA ^{ClaI (830)}
 841 AAAATTCGGT TAAGGCCAGG GGGAAAGAAG AAGTACAAGC TAAAGCACAT CGTATGGGCA AGCAGGGAGC
 911 TAGAACGATT CGCAGTTAAT CCTGGCCTGT TAGAAACATC AGAAGGCTGT AGACAAATAC TGGGACAGCT ^{AccI (959)}
 981 ACAACCATCC CTCAGACAG GATCAGAGGA GCTTCGATCA CTATACAACA CAGTAGCAAC CCTCTATTGT
 1051 GTGCACCAGC GGATCGAGAT CAAGGACACC AAGGAAGCTT TAGACAAGAT AGAGGAAGAG CAAAACAAGT
 1121 CCAAGAAGAA GGCCAGCAG GCAGCAGCTG ACACAGGACA CAGCAATCAG GTCAGCCAAA ATTACCCTAT

FIG. 11A

1191 AGTGCAGAAC ATCCAGGGGC AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAACGA TAAGCTTGGG
 1261 AGTTCCGCGT TACATAACTT ACGGTAAATG GCGCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC
 1331 GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGAGTAT
 1401 TTACGGTAAA CTGCCCACCTT GGCAGTACAT CAAGTGATC ATATGCCAAG TACGCCCCCT ATTGACGTCA
 1471 ATGACGGTAA ATGGCCCCGC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA
 1541 CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG
 1611 CGGTTTGACT CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG GAGTTTGTTT TGGCACCAAA
 1681 ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGGCGGTA GGCCTGTACG
 1751 GTGGGAGGTC TATATAAGCA GAGCTCGTTT AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT
 1821 TTTGACCTCC ATAGAAGACA CCGACTCTAG AGgatccATC TAAGTAAGCT TGGCATTCCG GTACTGTTGG
 1891 TAAAATGGAA GACGCCAAAA ACATAAAGAA AGGCCCGGCG CCATTCTATC CTCTAGAGGA TGAACCGCT
 1961 GGAGAGCAAC TGCATAAGGC TATGAAGAGA TACGCCCTGG TTCCTGGAAC AATTGCTTTT ACAGATGCAC
 2031 ATATCGAGGT GAACATCACG TACGCGGAAT ACTTCGAAAT GTCCGTTCCG TTGGCAGAAG CTATGAAACG
 2101 ATATGGGCTG AATACAAATC ACAGAATCGT CGTATGCAGT GAAACTCTC TTCAATTCTT TATGCCGGTG
 2171 TTGGCGCGCT TATTTATCGG AGTTGCAGTT GCGCCCGCGA ACGACATTTA TAATGAACGT GAATTGCTCA
 2241 ACAGTATGAA CATTTCGCAG CCTACCGTAG TGTTTGTTTC CAAAAGGGG TTGCAAAAAA TTTTGAACGT
 2311 GCAAAAAAAA TTACCAATAA TCCAGAAAAT TATTATCATG GATTCTAAAA CGGATTACCA GCGATTTCAG
 2381 TCGATGTACA CGTTCGTAC ATCTCATCTA CCTCCCGGTT TTAATGAATA CGATTTTGTA CCAGAGTCCT

FIG. 11B

2451 TTGATCGTGA CAAAACAATT GCACTGATAA TGAATTCCTC TGGATCTACT GGGTTACCTA AGGGTGTGGC

2521 CCTTCCGCAT AGAACTGCCT GCGTCAGATT CTCGCATGCC AGAGATCCTA TTTTGGCAA TCAAATCATT

2591 CCGGATACTG CGATTTTAAG TGTTGTTCCA TTCCATCAGC GTTTTGAAT GTTTACTACA CTCGGATATT

2661 TGATATGTGG ATTTCGAGTC GTCTTAATGT ATAGATTTGA AGAAGAGCTG TTTTACGAT CCCTTCAGGA

2731 TTACAAAATT CAAAGTCCGT TGCTAGTACC AACCCATT TTTCATTCTCG CCAAAGCAC TCTGATTGAC

2801 AAATACGATT TATCTAATT ACACGAAATT GCTTCTGGGG GCGCACCTCT TTCGAAAGAA GTCGGGGAAG

2871 CGGTTGCAAA ACGCTTCCAT CTTCAGGGA TACGACAAGG ATATGGGCTC ACTGAGACTA CATCAGCTAT

2941 TCTGATTACA CCCGAGGGGG ATGATAAACC GGGCGCGGTC GGTAAGTTG TTCCATTTTT TGAAGCGAAG

3011 GTTGTGGATC TGGATACCGG GAAAACGCTG GCGGTTAATC AGAGAGGCGA ATTATGTGTC AGAGGACCTA

3081 TGATTATGTC CGGTTATGTA AACAATCCGG AAGCGACCAA CGCCTTGATT GACAAGGATC GATGGCTACA

3151 TTCTGGAGAC ATAGCTTACT GGGACGAAGA CGAACACTTC TTCATAGTTG ACCGCTTGAA GTCTTTAATT

3221 AAATACAAAG GATATCAGGT GGCCCCGCT GAATTGGAAT CGATATTGTT ACAACACCCC AACATCTTCG

3291 ACGCGGGCGT GGCAGGTCTT CCCGACGATG ACGCCGGTGA ACTTCCCGCC GCCGTTGTTG TTTTGGAGCA

3361 CGGAAAGACG ATGACGGAAG AAGAGATCGT GGATTACGTC GCCAGTCAAG TAACAACCGC GAAAAAGTTG

3431 CGCGGAGGAG TTGTGTTTGT GGACGAAGTA CCGAAAGGTC TTACCGGAAA ACTCGACGCA AGAAAAATCA

3501 GAGAGATCCT CATAAAGGCC AAGAAGGGCG GAAAGTCCAA ATTGTAACTC GAGGGGGGGC CCGGTACCTT

3571 TAAGACCAAT GACTTACAAG GCAGCTGTAG ATCTTAGCCA CTTTTTAAAA GAAAAGGGGG GACTGGAAGG

3641 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

3711 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

3781 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

3851 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

3921 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

3991 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4061 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4131 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4201 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4271 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4341 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4411 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4481 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4551 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4621 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4691 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4761 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4831 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4901 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

4971 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5041 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5111 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5181 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5251 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5321 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5391 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5461 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5531 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5601 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5671 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5741 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5811 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5881 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

5951 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6021 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6091 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6161 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6231 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6301 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6371 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6441 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6511 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6581 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6651 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6721 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6791 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6861 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

6931 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7001 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7071 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7141 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7211 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7281 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7351 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7421 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7491 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7561 TGGTGGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT GCGGAGGAGT

7631 TGGTGG

FIG. 11C

3641 GCTAATTCAC TCCCAAAGAA GACAAGATAT CCTTGATCTG TGGATCTACC ACACACAAGG CTACTTCCCT
 3711 GATTGGCAGA ACTACACACC AGGGCCAGGG GTCAGATATC CACTGACCTT TGGATGGTGC TACAAGCTAG
 3781 TACCAGTTGA GCCAGATAAG GTAGAAGAGG CCAATAAAGG AGAGAACACC AGCTTGTTAC ACCCTGTGAG
 3851 CCTGCATGGA ATGGATGACC CTGAGAGAGA AGTGTTAGAG TGGAGGTTTG ACAGCCGCCT AGCATTTCAT
 3921 CACGTGGCCC GAGAGCTGCA TCCGGAGTAC TTCAAGAACT GCTGACATCG AGCTTGCTAC AAGGGACTTT
 3991 CCGCTGGGGA CTTTCCAGGG AGGCGTGGCC TGGGCGGGAC TGGGAGTGG CGAGCCCTCA GATGCTGCAT
 4061 ATAAGCAGCT GCTTTTGGCC TGTACTGGGT CTCTCTGGTT AGACCAGATC TGAGCCTGGG AGCTCTCTGG
 4131 CTAAGTAGGG AACCCACTGC TTAAGCCTCA ATAAAGCTTG CCTTGAGTGC TTCAAGTAGT GTGTGCCCGT
 4201 CTGTTGTGTG ACTCTGGTAA CTAGAGATCC CTCAGACCCT TTTAGTCAGT GTGGAAAATC TCTAGCACCC
 4271 CCCAGGAGGT AGAGGTTGCA GTGAGCCAAG ATCGCGCCAC TGCATTCCAG CCTGGGCAAG AAAACAAGAC
 4341 TGCTAAAAAT AATAATAATA AGTTAAGGGT ATTAATATATA TTTATACATG GAGGTCATAA AAATATATAT
 4411 ATTTGGGCTG GGCGCAGTGG CTCACACCTG CGCCCGGCC CTTGGGAGGC CGAGGCAGGT GGATCACCTG
 4481 AGTTTGGGAG TTCCAGACCA GCCTGACCAA CATGGAGAAA CCCCTTCTCT GTGTATTTTT AGTAGATTTT
 4551 ATTTTATGTG TATTTIATTC ACAGGTATTT CTGGAAAACT GAAACTGTTT TTCCTCTACT CTGATACCAC
 4621 AAGAATCATC AGCACAGAGG AAGACTTCTG TGATCAAAATG TGGTGGGAGA GGGAGGTTTT CACCAGCACA
 4691 TGAGCAGTCA GTTCTGCCGC AGACTCGCGG GGTGTCCTTC GGTTCAGTTC CAACACCGCC TGCCCTGGAGA
 4761 GAGGTCAGAC CACAGGCTGA GGGCTCAGTC CCCAAGACAT AAACACCCAA GACATAAACA CCCAACAGGT
 4831 CCACCCCGCC TGCTGCCCAG GCAGAGCCGA TTCACCAAGA CGGGAATTAG GATAGAGAAA GAGTAAAGTCA
 4901 CACAGAGCCG GCTGTGCGGG AGAACGGAGT TCTATTATGA CTCAAATCAG TCTCCCAAG CATTCGGGGA
 4971 TCAGAGTTTT TAAGGATAAC TTAGTGTGTA GGGGCCAGT GAGTTGGAGA TGAAAGCGTA GGGAGTCGAA
 5041 GCTGTCTTTT TGCGCCGAGT CAGTTCCTCG GTGGGGGCCA CAAGATCGGA TGAGCCAGTT TATCAATCCG
 5111 GGGGTGCCAG CTGATCCATG GAGTGCAGGG TCTGCAAAAT ATCTCAAGCA CTGATTGATC TTAGGTTTTA
 5181 CAATAGTGAT GTTACCCAG GAACAATTG GGGAAAGTCA GAATCTTGTA GCCTGTAGCT GCATGACTCC
 5251 TAAACCATAA TTTCTTTTT GTTTTTTTTT TTTTATTTTT GAGACAGGGT CTCACTCTGT CACCTAGGCT
 5321 GGAGTGCAGT GGTGCAATCA CAGCTCACTG CAGCCCCTAG AGCGGCCGCC ACCGCGGTGG AGCTCCAATT
 5391 CGCCCTATAG TGAGTCGTAT TACAATTCAC TGGCCGTCGT TTTACAACGT CGTGAAGTGG AAAACCCCTGG
 5461 CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC
 5531 ACCGATCGCC CTTCCCAACA GTTGCAGCAG CTGAATGGCG AATGGCGCGA AATTGTAAAC GTTAATATTT
 5601 TGTTAAAAAT CGCGTTAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT
 5671 CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGTTGAGT GTTGTTCAG TTTGGAACAA GAGTCCACTA
 5741 TTAAGAAGCG TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC
 5811 CATCACCTTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA AGCACTAAAT CGGAACCTTA AAGGGAGCCC
 5881 CCGATTTAGA GCTTGACGGG GAAAGCCGGG GAACGTGGCG AGAAAGGAA GGAAGAAAGC GAAAGGAGCG
 5951 GCGGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCA CCCCAGCGCG CTTAATCGCG
 6021 CGCTACAGGG CGCGTCCAG GTGGCACTTT TCGGGGAAAT GTGCGGGAA CCCCTATTTG TTTATTTTTT
 6091 TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA

FIG. 11D


```

6161 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT
6231 TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGCTTAC
6301 ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA
6371 GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCC
6441 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
6511 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC TCGCGCCAAC TTACTTCTGA
6581 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTTGA CAACATGGGG GATCATGTAA CTCGCCTTGA
6651 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG
6721 GCAACAACGT TCGCGAAACT ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT
6791 GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GCGCCTTCCG GCTGGCTGGT TTATTGCTGA
6861 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CCGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC
6931 CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA
7001 TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTGAGACCA AGTTTACTCA TATATACTTT AGATTGATTT
7071 AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT
7141 TAACGTGAGT TTTCTGTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
7211 TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
7281 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCTTT
7351 CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAAGTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA
7421 TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCTG TCTTACCGGG TTGGACTCAA GACGATAGTT
7491 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGTTTCG TGCACACAGC CCAGCTTGGG GCGAAGCACC
7561 TACACCGAAG TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG
7631 ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTCCAGGGG GAAACGCCTG
7701 GTATCTTTAT AGTCCTGTCT GGTTCGCCA CCTCTGACTT GAGCGTCCAT TTTTGTGATG CTCGTCAGGG
7771 GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCTTTT TACGGTTCCT GGCCTTTTGC TGGCCTTTTG
7841 CTCACATGTT CTTTCCTGCG TTATCCCTTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
7911 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCCAATA
7981 CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA CGACAGGTTT CCCGACTGGA
8051 AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTAGCT CACTCATTAG GCACCCCAAG CTTTACACTT
8121 TATGCTTCCG GCTCGTATGT TGTGTGGAAT TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC
8191 CATGATTACC CCAAGCTCGG AATTAAACCT CACTAAAGGG AACAAAAGCT GCTGCAGGGT CCCTAACTGC
8261 CAAGCCCCAC AGTGTGCCCT GAGGCTGCCC CTTCTTTCTA CCGGCTGCCC CCACTCGGCT TTGCTTTCCC
8331 TAGTTTCAGT TACTTGCGTT CAGCCAAGGT CTGAACTAG GTGCGCACAG AGCGGTAAGA CTGCGAGAGA
8401 AAGAGACCAG CTTTACAGGG GGTTTATCAC AGTGCACCCT GACAGTCGTC AGCCTCACAG GGGGTTTATC
8471 ACATTGCACC CTGACAGTCG TCAGCCTCAC AGGGGGTTTA TCACAGTGCA CCCTTACAAT CATTCATTT
8541 GATTACAAT TTTTTTAGTC TCTACTGTGC CTAAGTTGTA AGTTAAATTT GATCAGAGGT GTGTTCCAG
8611 AGGGGAAAAC AGTATATACA GGGTTCAGTA CTATCGCATT TCAGGCTTCC ACCTGGGTCT TGGAAATGTG
8681 CCCCCAGGG GTGATGACTA CCTCAGTTGG ATCTCCACAT GTCACAGTGA CACAAGATAA CCAAGACACC
8751 TCCCAAGGCT ACCACAATGG GCGGCCCTCC ACGTGACAT GGCCGGAGGA ACTGCCATGT CGGAGGTGCA
8821 AGCACACCTG CGCATCAGAG TCCTTGGTGT GGAGGGAGGG ACCAGCGCAG CTTCCAGCCA TCCACCTGAT
8891 GAACAGAACC TAGGGAAAGC CCCAGTTCTA CTTACACCAG GAAAGGC (SEQUENCE ID NO: 9)

```

FIG. 11E

mBCwCN frag	-----	-----	-----C-	--AC--G--	-----
m2BCwCN frag	-----	-----	-----C-	--G--G--	-----
BC/HXB2	-----	-----	-----	-----	-----
BC/NL43	-----	-----	-----	-----	-----
#1
	CGCGCACGGC	AAGAGGCGAG	GGGCGGCGAC	TGGTGAGTAC	GCCAAAAATT
mBCwCN frag	-----	-----	-----C-	C-----	-----
m2BCwCN frag	-----	-----	-----	-----	-----
BC/HXB2	-----T--	-----	-----	-----	-----
BC/NL43	-----	-----	-----	-----	-----G----
#51
	TTGACTAGCG	GAGGCTAGAA	GGAGAGAGAT	<u>GGGTGCGAGA</u>	GCGTCAGTAT
mBCwCN frag	-----	-----	-----	-----	-----
m2BCwCN frag	-----	-----	-----	-----	-----
BC/HXB2	-----	-----	-----	-----	-----
BC/NL43	-----	-----	AA-----	-----	-----
#101
	TAAGCGGGGG	AGAATTAGAT	CG		

FIG. 12

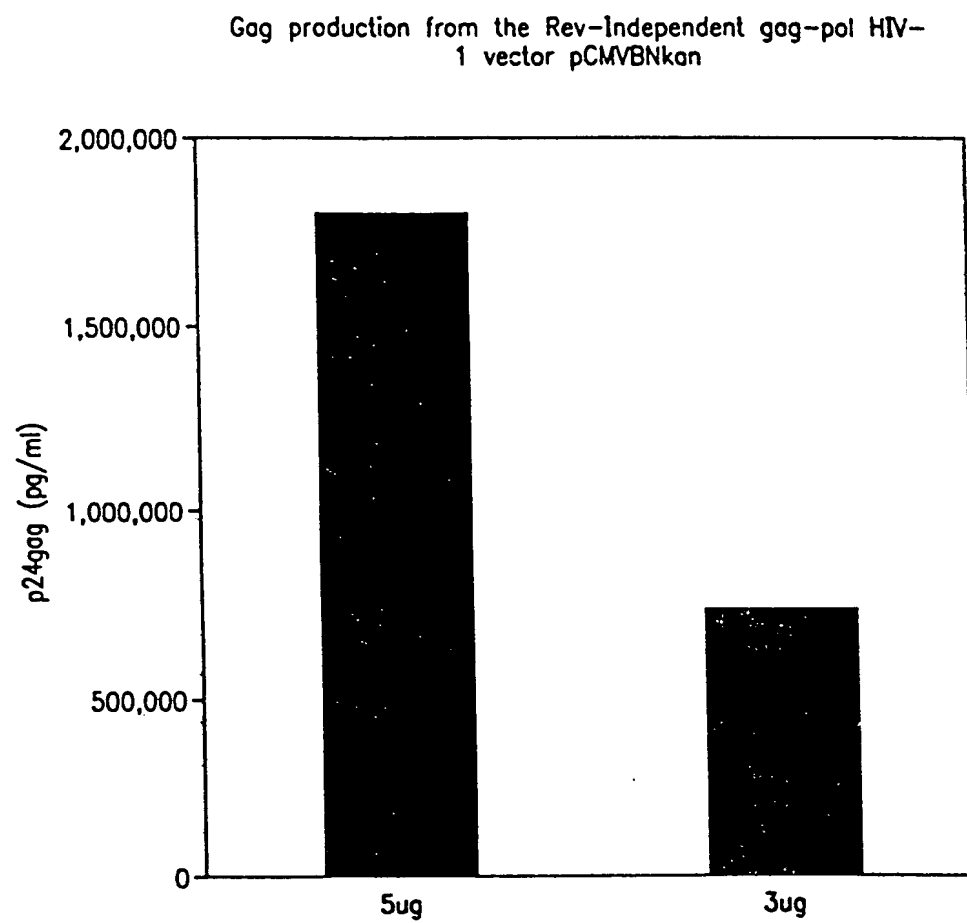


FIG. 13

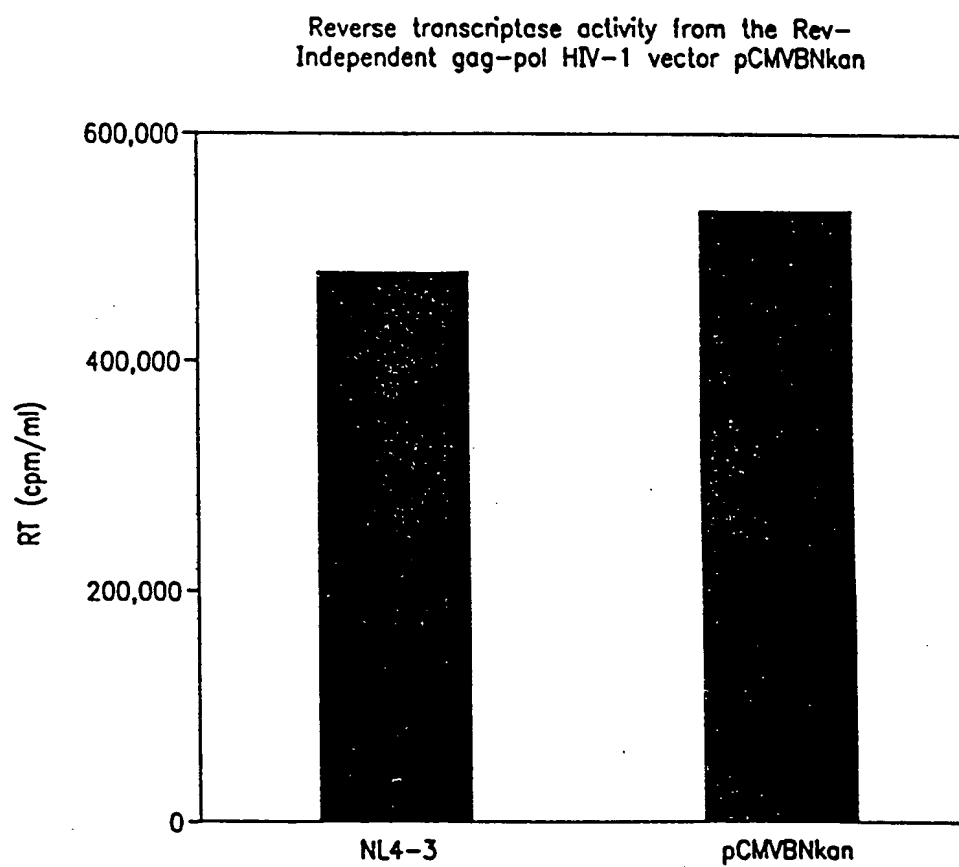


FIG. 14

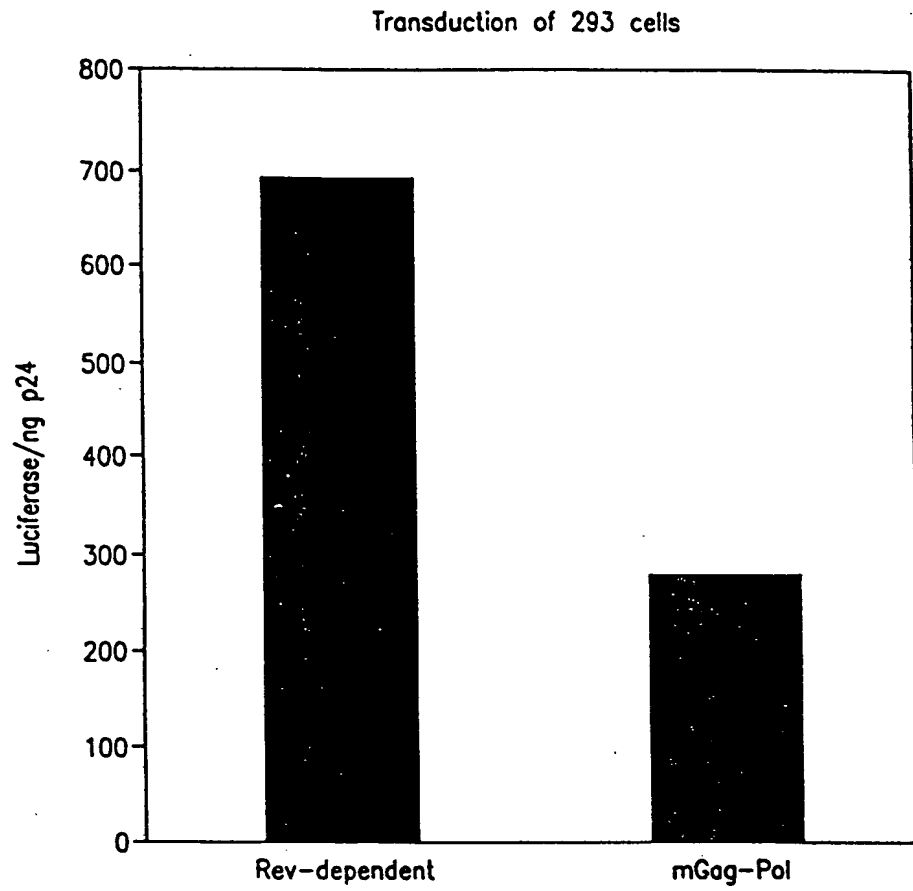


FIG. 15A

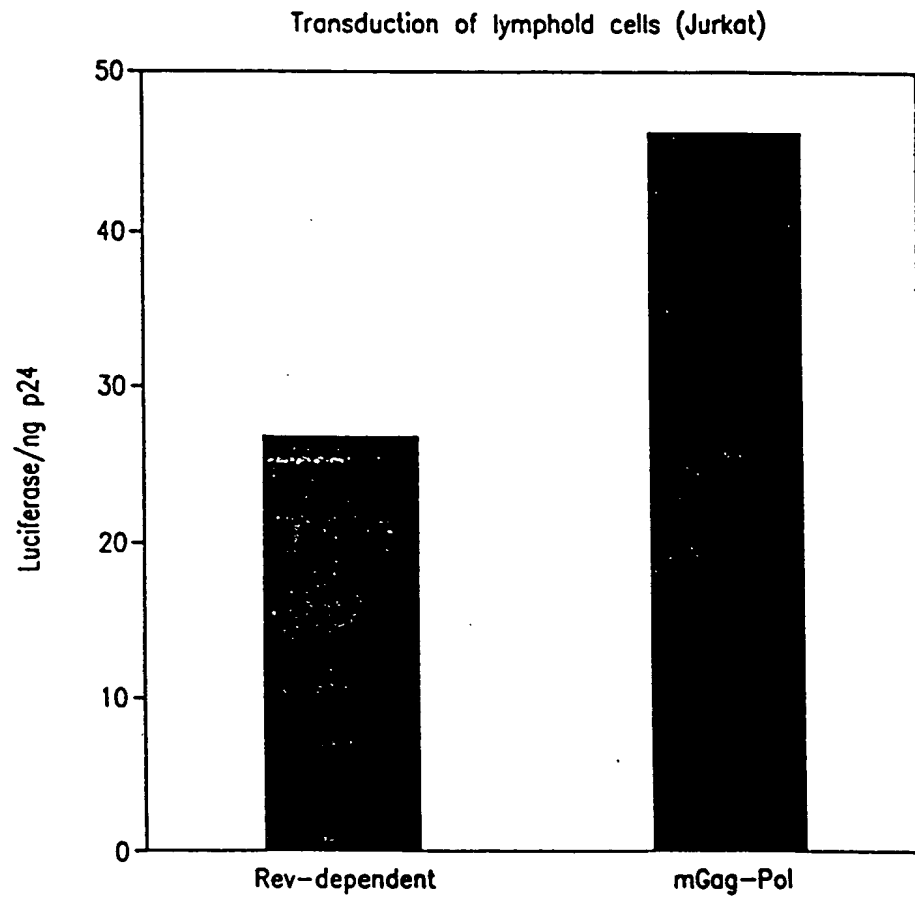


FIG. 15B

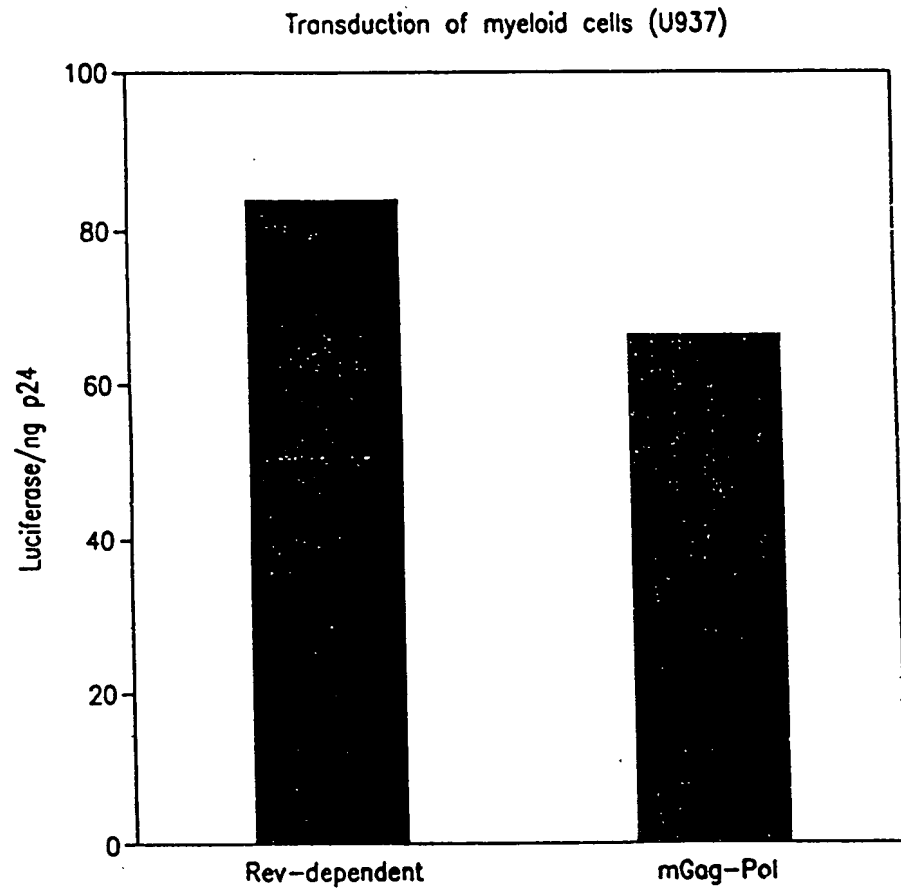


FIG. 15C

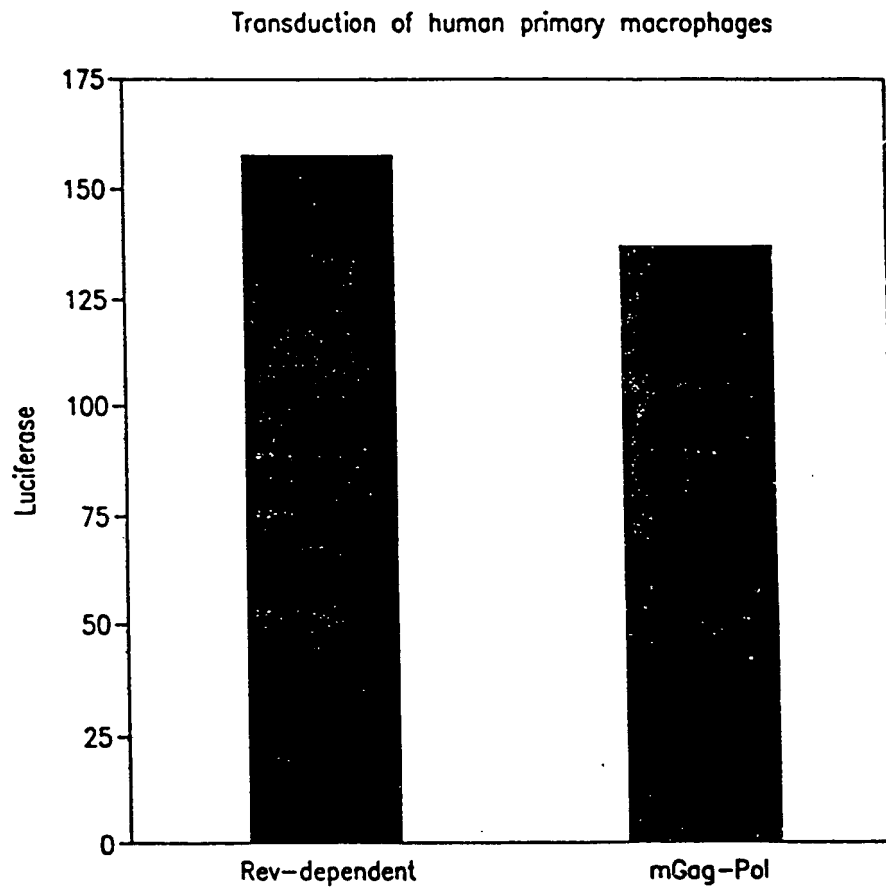


FIG. 15D

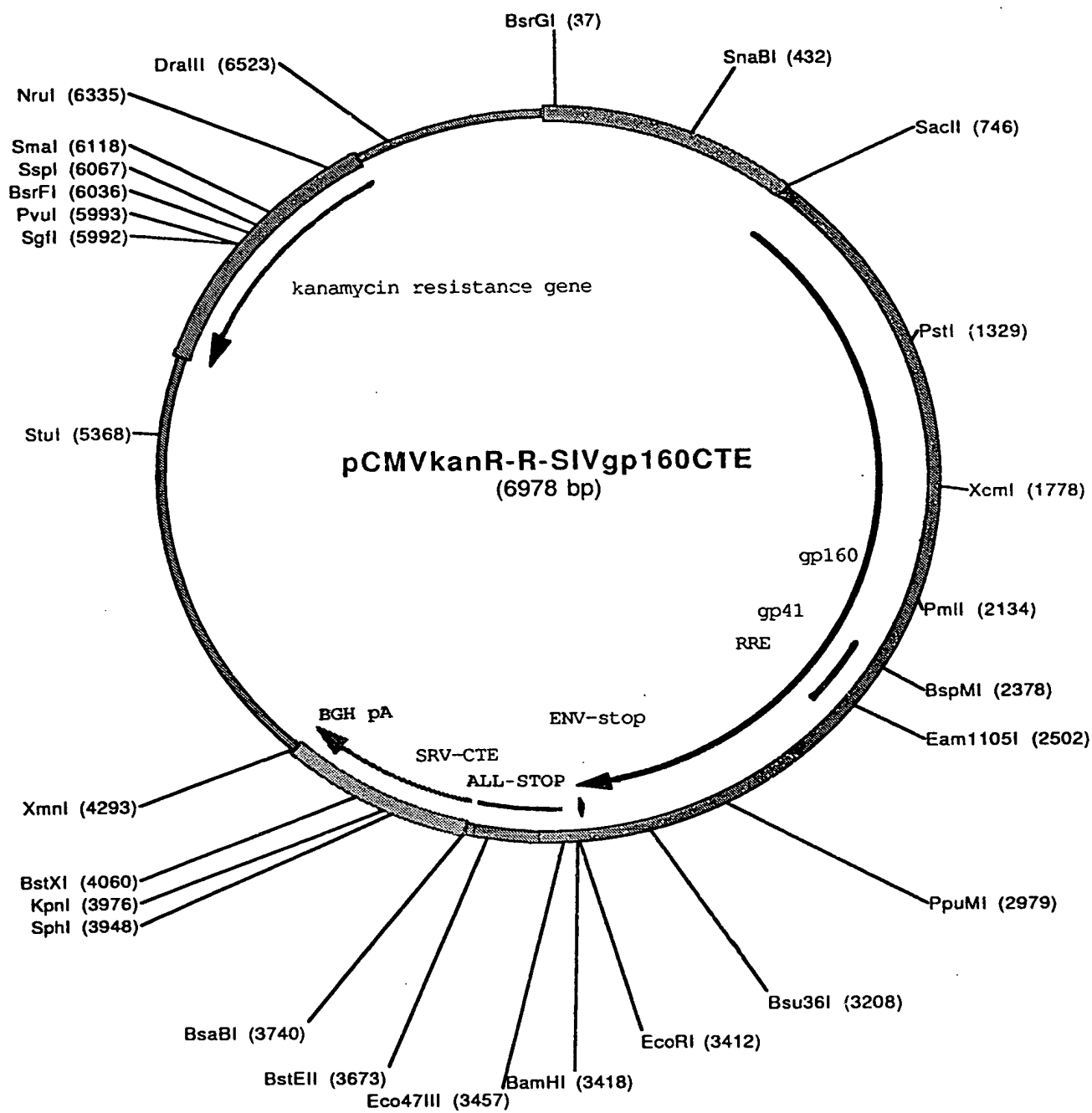


FIG. 16

42/45

BsrGI (37)
1 CCTGGCCATTGCATACGTTGTATCCATATCATAATGTACATTTATATGGCTCATGTCCAACATTACCGCCATGTTGA
81 CATTGATTATTGACTAGTTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTAC
161 ATAACCTACGGTAAATGGCCCGCCTGGCTGACCCCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCA
241 TAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
321 GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC

SnaBI (432)
401 CTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACA
481 TCAATGGGCGTGGATAGCGGTTTGAATCAGGGGATTTCGAAGTCTCAACCCATTGACGTCAATGGGAGTTTGTTTTGG
561 CACCAAAATCAACGGGACTTTCCAAAATGTCTGAACAACTCCGCCCAATTGACGCAAAATGGGCGGTAGGCGTGTACGGTG
641 GGAGGTCTATATAAGCAGAGCTCGTTTGTAGTGAACCGTCAGATCGCTGGAGACGCCATCCACGCTGTTTGGACCTCCATA

SacII (746)
721 GAAGACACCGGGACOGATCCAGCCTCCGCGGGCCGCGCTAAGTATGGGATGTCTTGGGAATCAGCTGCTTATCGCCATCT

1►Met Gl yCysLeuGl yAsnGl nLeuLeuIleAl alleL
801 TGCTTTTAAGTGTCTATGGGATCTATTGTACTCTATATGTCACAGTCTTTTATGGTGTACCAGCTTGGAGGAATGCGACA

13►euLeuLeuSer Val TyrGl yIleTyrCysThr LeuTyrVal Thr Val PheTyrGl yVal ProAl aT rpArgAsnAl aThr
881 ATTCCCTCTTTTGTGCAACCAAGAATAGGGATACTTGGGGAACAACTCAGTGCCTACCAGATAATGGTGATTATTGAGA

40►IleProLeuPheCysAl aThr LysAsnArgAspThr TrpGl yThr Thr Gl nCysLeuProAspAsnGl yAspTyrSer Gl
961 AGTGGCCCTTAATGTTACAGAAAGCTTTGATGCCTGGAATAATACAGTCACAGAACAGGCAATAGAGGATGTATGGCAAC

66►uValAl aLeuAsnVal Thr Gl uSer PheAspAl aT rpAsnAsnThr Val Thr Gl uGl nAl alleGl uAspVal TrpGl nL
1041 TCTTTGAGACCTCAATAAAGCCTTGTGTAATAATTATCCCCATTATGCATTACTATGAGATGCAATAAAGTGAGACAGAT

93►euPheGl uThr Ser IleLysProCysVal LysLeuSer ProLeuCysIleThrMetArgCysAsnLysSer Gl uThrAsp
1121 AGATGGGATTGACAAAATCAATAACAACACAGCATCAACAACATCAACGACAGCATCAGCAAAAGTAGACATGGTCAA

120►A rgTrpGl yLeuThr LysSer IleThr Thr Thr Al aSer Thr Thr Ser Thr Thr Al aSer Al aLysVal AspMet ValAs
1201 TGAGACTAGTTCTTGTATAGCCAGGATAATTGCACAGGCTTGGACAAGAGCAAAATGATAAGCTGTAAATCAACATGA

146►nGl uThr Ser Ser CysIleAl aGl nAspAsnCysThr Gl yLeuGl uGl nGl uGl nMet IleSer CysLysPheAsnMet T
PstI (1329)
1281 CAGGGTTAAAAAGAGACAAGAAAAAGAGTACAATGAACTTGGTACTCTGCAGATTGGTATGTGAACAAGGGAATAAC

173►hr Gl yLeuLysArgAspLysLysLysGl uTyrAsnGl uThr TrpTyrSer Al aAspLeuVal CysGl uGl nGl yAsnAsn
1361 ACTGGTAATGAAAGTAGATGTTACATGAACCACTGTAACTTCTGTTATCCAAGAGTCTGTGTGACAAACATTATTGGGA

200►Thr Gl yAsnGl uSer ArgCysTyrMetAsnHisCysAsnThr Ser Val IleGl nGl uSer CysAspLysHisTyrTrpAs
1441 TGCTATTAGATTAGGTATTGTGCACCTCCAGGTTATGCTTTGCTTAGATGTAATGACACAAATTATTAGGCTTTATGC

226►pAl alleArgPheArgTyrCysAl aProProGl yTyrAl aLeuLeuArgCysAsnAspThrAsnTyrSer Gl yPheMet P
1521 CTAAATGTTCTAAGGTGGTGTCTCTCATGCACAAGGATGATGGAGACACAGACTTCTACTTGGTTGGCTTTAATGGA

253►roLysCysSer LysVal Val Val Ser Ser CysThr ArgMetMet Gl uThr Gl nThr Ser Thr TrpPheGl yPheAsnGl y
1601 ACTAGAGCAGAAAATAGAATTATATTACTGGCATGGTAGGATAATAGGACTATAATTAGTTTAAATAAGTATTATAA

280►Thr ArgAl aGl uAsnArgThr TyrIleTyrTrpHisGl yA rgAspAsnArgThr IleIleSer LeuAsnLysTyrTyrAs
1681 TCTAACAAATGAAATGTAGAAGACCAGGAAATAAGACAGTTTTTACCAGTCACCATTATGTCTGGATTGGTTTTCCACTCAC

306►nLeuThrMetLysCysArgArgProGl yAsnLysThr Val LeuProVal Thr IleMetSer Gl yLeuVal PheHisSer G
XcmI (1778)
1761 AACCAATCAATGATAGCCAAAGCAGGCATGGTGTGGTTTGGAGGAAATGGAAGGATGCAATAAAGAGGTGAAGCAG

333►InProIleAsnAspArgProLysGl nAl aT rpCysTrpPheGl yGl yLysTrpLysAspAl alleLysGl uVal LysGl n
1841 ACCATTGTCAAACATCCAGGTATACTGGAATAACAATACTGATAAAATCAATTTGACGGCTCTGGAGGAGGAGATCC

360►Thr IleVal LysHisProArgTyrThr Gl yThrAsnAsnThrAspLysIleAsnLeuThr Al aProGl yGl yGl yAspPr
1921 GGAAGTTACCTTCATGTGGACAAATTGCAGAGGAGAGTTTCTCTACTGTAATGAATTGGTTTCTAAATTGGGTAGGAAG

386►oGl uVal Thr PheMetTrpThrAsnCysArgGl yGl uPheLeuTyrCysLysMetAsnTrpPheLeuAsnTrpVal Gl uA
2001 ATAGGAATACAGCTAACCAAGCCAAAGGAACAGCATAAAGGAATTAAGTGCCATGTATATTAGACAAATAATCAAC

413►spArgAsnThr Al aAsnGl nLysProLysGl uGl nHisLysArgAsnTyrVal ProCysHisIle ArgGl nIleIleAsn

FIG. 17

2081 ACTTGGCATAAAGTAGGCAAAAATGTTTATTTGCTCCAAGAGAGGGAGACCTCACGTGTAACCTCCACAGTGACCAGTCT

PmlI (2134)

440 Thr TrpHisLysVal Gl yLysAsnVal TyrLeuProProArgGl uGl yAspLeuThr CysAsnSer Thr Val Thr Ser Le
2161 CATAGCAAACATAGATTGGATTGATGGAAACCAACTAATATCACCATGAGTGCAGAGGTGGCAGAACTGTATCGATTGG

466 u l l eAl aAsn l l eAspTrp l l eAspGl yAsnGl nThrAsn l l eThrMetSerAl aGl uValAl aGl uLeuTyrArgLeuG
2241 AATTGGGAGATTATAAATTAGTAGAGATCACTCCAATTGGCTTGGCCCCACAGATGTGAAGAGGTACACTACTGTTGGC

493 l uLeuGl yAspTyrLysLeuVal Gl u l l eThrPro l l eGl yLeuAl aProThrAspVal LysArgTyrThr Thr Gl yGl y
BspMI (2378)

2321 ACCTCAAGAAATAAAGAGGGGTCTTTGTGCTAGGGTTCTTGGGTTTTCTCGCAACGGCAGGTTCTGCAATGGAGCGCG

520 Thr Ser ArgAsnLysArgGl yVal PheVal LeuGl yPheLeuGl yPheLeuAl aThrAl aGl ySerAl aMetGl yAl aAl
2401 CAGCCTGACCTCACGGCACAGTCCCGAACTTTATTGGCTGGGATAGTCCAACAGCAGCAACAGCTGTTGGACGTGGTCA

546 aSerLeuThrLeuThrAl aGl nSerArgThrLeuLeuAl aGl y l l eVal Gl nGl nGl nGl nGl nLeuLeuAspValValL
Eam1105I (2502)

2481 AGAGACAACAAGAATTGTTGCGACTGACCGTCTGGGGAACAAGAACCCTCCAGACTAGGGTCACTGCCATCGAGAAGTAC

573 ysArgGl nGl nGl uLeuLeuArgLeuThrValTrpGl yThrLysAsnLeuGl nThrArgValThrAl a l l eGl uLysTyr
2561 TTAAAGGACCAAGGCGCAGCTGAATGCTTGGGGATGTGCGTTTAGACAAGTCTGCCACACTACTGTACCATGGCCAAATGC

600 LeuLysAspGl nAl aGl nLeuAsnAl aTrpGl yCysAl aPheArgGl nValCysHisThrThrValProTrpProAsnAl
2641 AAGTCTAACACCAAGTGGACAATGAGACTTGGCAAGAGTGGGAGCGAAAGTTGACTTCTTGAAGAAAATATAACAG

626 aSerLeuThrProLysTrpAsnAsnGl uThrTrpGl nGl uTrpGl uArgLysValAspPheLeuGl uGl uAsn l l eThrA
2721 CCCTCTAGAGGAGGCACAAATTCACAAGAGAAGAATATGAATTACAAAGTTGAATAGCTGGGATGTGTTTGGC

653 l aLeuLeuGl uGl uAl aGl n l l eGl nGl nGl uLysAsnMetTyrGl uLeuGl nLysLeuAsnSerTrpAspValPheGl y
2801 AATTGGTTTGACCTTGCTTCTTGATAAAGTATATACAATATGGAGTTTATATAGTTGTAGGAGTAATACTGTTAAGAAT

680 AsnTrpPheAspLeuAl aSerTrp l l eLysTyr l l eGl nTyrGl yValTyr l l eValValGl yVal l l eLeuLeuArg l l
2881 AGTGATCTATATAGTACAAATGCTAGCTAAGTTAAGGCAGGGTATAGGCCAGTGTCTCTTCCCAACCTCTTATTTC

706 eVal l l eTyr l l eVal Gl nMetLeuAl aLysLeuArgGl nGl yTyrArgProValPheSerSerProProSerTyrPheG
PpuMI (2979)

2961 AGCAGACCCATATCCAACAGGACCCGCACTGCCAACAGAGAAGGCAAGAAAGAGACGGTGGAGAAGCGGTGGCAAC

733 l nGl nThrHis l l eGl nGl nAspProAl aLeuProThrArgGl uGl yLysGl uArgAspGl yGl yGl uGl yGl yAsn
3041 AGCTCTGGCTTGGCAGATAGAATATATCCACTTTCTTATTGTCAGCTTATTAGACTCTTGACTTGGCTATTTCAGTAA

760 SerSerTrpProTrpGl n l l eGl uTyr l l eHisPheLeu l l eArgGl nLeu l l eArgLeuLeuThrTrpLeuPheSerAs
3121 CTGTAGGACTTTGCTATCGAGAGTATACCATCTCCAACCAATACTCCAGAGGCTCTCTCGACCTACAGAGGATTC

786 nCysArgThrLeuLeuSerArgValTyrGl n l l eLeuGl nPro l l eLeuGl nArgLeuSerAl aThrLeuGl nArg l l eA
Bsu36I (3208)

3201 GAGAAGTCTCAGGACTGAAGTACCTACCTACAATATGGTGGAGCTATTTCCATGAGGCGGTCCAGGCCGTCTGGAGA

813 rgGl uValLeuArgThrGl uLeuThrTyrLeuGl nTyrGl yTrpSerTyrPheHisGl uAl aVal Gl nAl aValTrpArg
3281 TCTGCGACAGAGACTTGTGGGGCGGTGGGAGACTTATGGGAGACTCTTAGGAGAGGTGGAAGATGGATACTCGCAAT

840 SerAl aThrGl uThrLeuAl aGl yAl aTrpGl yAspLeuTrpGl uThrLeuArgArgGl yGl yArgTrp l l eLeuAl a l l
BamHI (3418)

EcoRI (3412)

3361 CCCAGGAGGATTAGACAAGGGCTTGAGCTCACTCTCTTGTGAGGGACAGAGAATTCGGATCCactagtctctagaCTOGA

866 eProArgArg l l eArgGl nGl yLeuGl uLeuThrLeuLeu* * *

Eco47III (3457)

3441 GGGGGGGCCCGGTACGAGCGCTTAGCTAGCTAGAGACCCTCCCTGCGAGCTAAGCTGGACAGCCAATGACGGGTAAG

3521 AGAGTGACATTTTCACTAACCTAAGACAGGAGGGCGTCAGAGCTACTGCCTAATCCAAGACGGGTAAAAGTGATAAA

BstEII (3673)

3601 AATGTATCACTCCAACCTAAGACAGGCGCAGCTTCCGAGGGATTTGTCGTCTGTTTATATATATTAAAGGGTGACCT

3681 GTCCGGAGCCGTGCTGCCCGGATGATGTCTTGGTCTAGACTCGAGGGGGGGCCCGGTACGATOCAGATCTGTGTGCTT
 3761 CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCGGTGCTTCTTGACCTGGAAGGTGCCACTCCCACTGTCCTTTCC
 3841 TAATAAAATGAGGAAATGCATCGCATTGTCTGAGTAGGTGTCTATTCTATTCTGGGGGTGGGGTGGGACAGCACAGCAA
 3921 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTATGGGTACCGAGGTGCTGAAGAATTGAC
 4001 CCGGTTCTCTCTGGGCCAGAAAGAAGCAGGCACATCCCTTCTCTGTGACACACCTGTCCACGCCCTGGTTCTTAGTT
 4081 CCAGCCCCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCAATCCACCGCTAAAGTACTTGGAGCGGTCTC
 4161 TCCCTCCCTCATCAGCCACCAACCAACCTAGCCTCCAAGAGTGGGAAGAAATTAAAGCAAGATAGGCTATTAAGTGC
 4241 AGAGGGAGAGAAAATGCCCTCAACATGTGAGGAAGTAATGAGAGAAATCATAGAATTTCTTCGCTTCTCGCTCACTGA
 4321 CTCGCTGCGCTCGGTCTGCTGCGCTGCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAATACGGTTATCCACAGAATCAG
 4401 GGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAGGCCCGCGTTGCTGGCGTTT
 4481 TTCCATAGGCTCCGCCCCCTGAAGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATA
 4561 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCG
 4641 CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAAGCTGTAGGTATCTCAGTTCCGTTGAGGTGCTTCGCTCC
 4721 AAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAACCC
 4801 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG
 4881 AGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
 4961 TTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCA
 5041 GATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAACT
 5121 CACGTTAAGGATTTTGGTCAATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAAAAATGAAGTTTAAA
 5201 TCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
 5281 TCTATTTCGTTTCATCCATAGTTGCCTGACTCCGGGGGGGGGGCGCTGAGGTCTGCCTCGTGAAGAAGGTGTTGCTGAC
 5361 TCATACCAGGCTGAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCAGGTTGATGAGAGCTTTGTTGTAGGTGGA
 5441 CCAGTTGGTGATTTTGAACITTTGCTTTGCCACGGAACGGTCTGCGTTGTGCGGAAGATGCGTGATCTGATCCTTCAACT
 5521 CAGCAAAAGTTGATTTATTCAACAAAGCCCGCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAA
 5601 CCAATTCGTATTAGAAAACTCATCGAGCATCAAAATGAAGTCAATTTTATCATATCAGGATTATCAATACCAATATTT
 2714 PhePheGluAspLeuMetLeuHisPheGlnLeuLysAsnMetAspProAsnAspIleGlyTyrLysG
 5681 TGAAAAAGCCGTTTCTGTAATGAAGGAGAAACTCACGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGC
 2484 InPheLeuArgLysGlnLeuSerProSerPheGluGlyLeuCysAsnTrpLeuIleAlaLeuAspGlnTyrArgAspAla
 5761 GATTCGAGCTCGTCCAAATCAATCAACCTATTAAATTTCCCGCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCAT
 2224 IleGlyValAlaGlyValAspIleCysGlyIleLeuLysGlyGluAspPheIleLeuAsnAspLeuSerPheAspGlyHis
 5841 GAGTGACGACTGAATCCGGTGAGAAATGGCAAGCTTATGCATTTCTTCCAGACTGTTCACAGGCCAGCCATTACGC
 1954 sThrValValSerAspProSerPheProLeuLeuLysHisMetGluLysTrpValGlnGluValProTrpGlyAsnArgG
 5921 TCGTCATCAAAATCACTCGCATCAACCAACCGTTATTCATTCGTGATTGCGCTGAGCGAGACGAAATACGGATCGCT
 1684 LuAspAspPheAspSerAlaAspValLeuGlyAsnAsnMetArgSerGlnAlaGlnAlaLeuArgPheValAlaArgAspSer
 6001 GTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAGCGCATCAACAAATATTTTACCTG
 1424 AsnPheProCysAsnCysValProIleSerHisLeuArgArgLeuPheValAlaLeuAlaAspValIleAsnGluGlySe
 6081 AATCAGGATATCTTCTAATACTGGAATGCTGTTTTCCCGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTA
 1154 rAspProTyrGluGluLeuValGlnPheAlaThrLysGlyProIleAlaThrThrLeuLeuTrpAlaAspAspProThrA
 6161 CGGATAAAATGCTTGATGGTGGGAAGAGGCATAAATCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATT
 884 rGllePheHisLysIleThrProLeuProMetPheGluThrLeuTrpAsnLeuArgValMetGluAspThrValAspAsn
 6241 GGCAACGCTACCTTTGCCATGTTTTCAGAAACAACTCGGCGCATCGGGCTTCCCATACAATCGATAGATTGTGCGACCTG
 624AlaValSerGlyLysGlyHisLysLeuPheLeuGluProAlaAspProLysGlyTyrLeuArgTyrIleThrAlaGlySe
 6321 ATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAA
 354 rGlnGlyValAsnAspArgAlaTrpLysTyrGlyTyrLeuAspAlaAspMetAsnSerAsnLeuArgProArgSerCysS
 6401 GACGTTCCCGTTGAATATGGCTCATACACCCCTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTTCATGATGA
 84 rThrGluArgGlnIleHisSerMet
 6481 TATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCCCCCCCCCCCCATTATTGAAGCA
 5661 TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCAC

45/45

6641 TTTCCCGAAAAGTGCCACCTGACGTCTAAGAAAACATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAG
6721 GCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTC
6801 TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACAT
6881 GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCG
6961 CATCAGATTGGCTATTGG